UBLICATIONS DE L'INSTITUT DE STATISTIQUE DE L'UNIVERSITÉ DE PARIS

MÉMOIRES ET CONFÉRENCES SUR LE CALCUL DES PROBABILITÉS, LA STATISTIQUE THÉORIQUE ET APPLIQUÉE, L'ÉCONOMÉTRIE

VOL. IX — FASCICULE 4 - 1960

PARIS

11, Rue Pierre-Curie

′UBLICATIONS DE L'INSTITUT DE STATISTIQUE
DE L'UNIVERSITÉ DE PARIS

MÉMOIRES ET CONFÉRENCES SUR LE CALCUL DES PROBABILITÉS,
LA STATISTIQUE THÉORIQUE ET APPLIQUÉE, L'ÉCONOMÉTRIE

VOL. IX – FASCICULE 4 - 1960

PARIS

11, Rue Pierre-Curie

Toute la correspondance relative aux publications
doit être envoyée à l'adresse

INSTITUT DE STATISTIQUE DE L'UNIVERSITÉ DE PARIS
Institut Henri Poincaré - 11, Rue Pierre Curie - Paris (5e)

Les manuscrits doivent être envoyés à M. Daniel DUGUE
à l'adresse précédente.

CALCUL D'UNE INTÉGRALE AU MOYEN DE LA SUITE $X_n = A_n$. ÉVALUATION DE L'ERREUR

Jean Paul BERTRANDIAS

Si on utilise la méthode de Monte-Carlo pour calculer l'intégrale

$$J = \int_0^1 f(x)\, dx \qquad (1)$$

on est conduit à étudier la limite lorsque N augmente indéfiniment de

$$J_N = \frac{1}{N} \sum_{n=0}^{N-1} f(x_n) \qquad (2)$$

où x_n est une suite de "nombres aléatoires".

On obtient une convergence de J_N vers J en $\frac{1}{\sqrt{N}}$, la convergence étant prise, bien entendu, du sens du calcul des probabilités.

Dans un article paru précédemment, M. Bass [1] utilise dans (2) une suite de nombres bien déterminés par des procédés arithmétiques et la convergence de J_N est donnée par les théorèmes de H. Weyl [4].

M. Bass étudie plus particulièrement le cas de la suite $\underset{\smile}{An}$ (parties fractionnaires des multiples successifs d'un nombre réel A). Nous nous proposons de préciser ses conclusions et de montrer que dans des cas assez généraux, l'erreur commise est inférieure à $\frac{S}{N}$, S étant un nombre que l'on peut calculer numériquement lorsque A admet un développement en fraction continue simple.

Ce travail a été effectué avec l'aide de M. Guilloud pour la partie numérique. De nombreux calculs de moyennes ont été faits sur machine à calculer électronique et quelques unes des courbes les plus caractéristiques sont reproduites et commentées ici.

I - MAJORATION DE LA DIFFERENCE $|J - J_N|$ -

Le développement de f(x) en série de Fourier dans l'intervalle (0,1) est de la forme :

$$\sum_{k=-\infty}^{+\infty} c_k e^{2i\pi kx} \quad \text{avec} \quad c_{-k} = \bar{c}_k \tag{3}$$

On suppose qu'il existe deux nombres strictement positifs M et λ tels que la série c_k admette comme majorante la série

$$\gamma_k = \frac{M}{|k|^{1+\lambda}} \tag{4}$$

La différence entre l'intégrale J et sa valeur approchée J_N est :

$$I_N = \sum_{\substack{k=-\infty \\ k\neq 0}}^{+\infty} c_k \frac{1}{N} \sum_{n=0}^{N-1} e^{2i\pi kAn} = \frac{1}{N} \sum_{k\neq 0} c_k \frac{e^{2i\pi kAN} - 1}{e^{2i\pi kA} - 1} \tag{5}$$

$$= \frac{1}{N} \sum_{k\neq 0} c_k \frac{\sin \pi kAN}{\sin \pi kA} \; e^{i\pi kA(N-1)} \tag{6}$$

On peut majorer $|I_N|$ par $\frac{2}{N}$ $\mathcal{S}$ avec

$$\mathcal{S} = \sum_{k=1}^{\infty} \frac{|c_k|}{|\sin \pi kA|} \tag{7}$$

En tenant compte de (4) et de l'inégalité

$$|\sin \pi kA| \geqslant 2 \, \|kA\| \tag{8}$$

où $\|kA\|$ = minimum de $\underset{\smile}{kA}$ ou $1 - \underset{\smile}{kA}$

on obtient .

$$|I_N| < \frac{1}{N} \sum_{k=1}^{\infty} \frac{\gamma_k}{\|kA\|} \tag{9}$$

On peut enfin appliquer à cette majoration la transformation d'Abel :

$$|I_N| < \frac{1}{N} \sum_{k=1}^{\infty} (\gamma_j - \gamma_{j+1}) \, \sigma(j) = \frac{S}{N} \tag{10}$$

avec

$$\sigma(j) = \sum_{k=1}^{j} \frac{1}{\| kA \|} \qquad (11)$$

Nous nous proposons de montrer que lorsque A est un nombre algébrique, I_N est de l'ordre de $\frac{1}{N}$, c'est-à-dire que la série S intervenant au second membre de (10) est convergente. Pour cela, il suffit de trouver une borne supérieure assez précise de $\sigma(j)$ en fonction de j.

II - ETUDE DE σ(j) -

Nous utiliserons le développement de A en fraction continue, Rappelons d'abord quelques définitions et quelques propriétés classiques.

Les quotients incomplets a_i sont définis par

$$\begin{array}{lll} A = \hat{A} + \frac{1}{u_1}, & & u_1 > 1 \\ u_1 = a_1 + \frac{1}{u_2}, & a_1 = \hat{u}_1, & u_2 > 1 \\ \text{--------------------} & & \\ u_i = a_i + \frac{1}{u_{i+1}}, & a_i = \hat{u}_i, & u_{i+1} > 1 \\ \text{--------------------} & & \end{array} \qquad (12)$$

le symbole $\hat{u}$ désignant la partie entière du nombre u. Si A est irrationnel, u_i existe toujours et $a_i \geqslant 1$ quel que soit i. L'ensemble des a_i définit ce qu'on appelle le développement de A en fraction continue.

La réduite d'ordre i est la fraction $\frac{P_i}{Q_i}$ obtenue en remplaçant u_i par sa partie entière :

$$\frac{P_i}{Q_i} = \hat{A} + \cfrac{1}{\hat{u}_1 + \cfrac{1}{\hat{u}_2 \dots + \cfrac{1}{\hat{u}_i}}} \qquad (13)$$

Elle représente une approximation de A par un nombre rationnel.

Les P_i et les Q_i vérifient les relations suivantes :

$$Q_l = a_l Q_{l-1} + Q_{l-2}$$

$$P_l = a_l P_{l-1} + P_{l-2} \qquad (14)$$

avec $\qquad Q_0 = 1 \qquad Q_1 = a_1$

et $\qquad P_0 = a_0 = \hat{A} \qquad P_1 = a_0 a_1 + 1$

En éliminant a_l de ces relations :

$$P_l Q_{l-1} - Q_l P_{l-1} = -(P_{l-1} Q_{l-2} - P_{l-2} Q_{l-1})$$

$$= (-1)^{l-1} [P_1 Q_0 - P_0 Q_1] = (-1)^{l-1}$$

D'où :

$$\frac{P_l}{Q_l} - \frac{P_{l-1}}{Q_{l-1}} = \frac{(-1)^{l-1}}{Q_l Q_{l-1}}$$

Puis, comme A est compris entre deux réduites successives :

$$\lim \frac{P_l}{Q_l} = A = \frac{P_0}{Q_0} + \sum_{n=1}^{\infty} \left(\frac{P_n}{Q_n} - \frac{P_{n-1}}{Q_{n-1}} \right)$$

$$A = \hat{A} + \frac{1}{Q_0 Q_1} - \frac{1}{Q_1 Q_2} + \ldots \qquad (15)$$

Les autres propriétés des fractions continues sur lesquelles nous nous appuierons sont les suivantes :

Propriété I - Pour tous les entiers positifs k inférieurs à $Q_{\nu+1}$, on a :

$$\| kA \| \geqslant \| Q_\nu A \| = \delta_\nu \qquad (16)$$

$$\delta_\nu = \frac{1}{Q_{\nu+1}} - \frac{Q_\nu}{Q_{\nu+1} Q_{\nu+2}} + \frac{Q_\nu}{Q_{\nu+2} Q_{\nu+3}} - \ldots \qquad (17)$$

l'égalité n'ayant lieu que pour $k = Q_\nu$.

Propriété II - Les Q_μ points d'affixes

$$e^{2i\pi(m-Q_\mu+1)A}, \; e^{2i\pi(m-Q_\mu+2)A}, \ldots, \; e^{2i\pi m A}$$

$$(m < Q_{\nu+1}, \quad \mu \leqslant \nu)$$

sont, sur la circonférence $|z| = 1$, les sommets d'un polygône convexe P_μ^ν de Q_μ cotés. Les arcs sous-tendus par les plus petits cotés sont égaux à $2\pi\delta_{\mu-1}$.

La propriété I est une propriété classique des fractions continues exprimant que les $\frac{P_l}{Q_l}$ sont d'une certaine façon les "meilleures" approximations rationnelles de A (voir par exemple [3] théorème 182). La formule (17) se démontre facilement à partir de (15) et la propriété II est une conséquence immédiate de I.

D'autre part si $Q_\nu \leqslant j < Q_{\nu+1}$, on peut écrire :

$$j = b_\nu Q_\nu + b_{\nu-1} Q_{\nu-1} + \dots + b_0 Q_0 \tag{18}$$

avec

$$0 \leqslant b_\mu \leqslant a_{\mu+1} \tag{19}$$

Alors la sommation (11) donnant $\sigma(j)$ peut se décomposer :

$$\sigma(j) = \sum_{k=1}^{Q_\nu} \frac{1}{\|kA\|} + \sum_{Q_\nu+1}^{2Q_\nu} + \dots + \sum_{(b_\nu-1)Q_\nu+1}^{b_\nu Q_\nu}$$

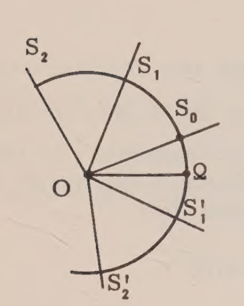
$$+ \sum_{b_\nu Q_\nu+1}^{b_\nu Q_\nu+Q_{\nu-1}} + \dots + \dots + \sum_{b_\nu Q_\nu+\dots+(b_0-1)Q_0+1}^{j} \tag{20}$$

Chaque signe $\sum$ porte sur Q_μ multiples consécutifs de A parmi les j premiers avec $\mu \leqslant \nu$ et $j < Q_{\nu+1}$. On peut donc associer à chaque somme de (20) un des polygônes P_μ^ν défini plus haut et on notera la somme correspondante par $\sum_{P_\mu^\nu}$. Soit alors M_μ^ν une borne supérieure de $\sum_{P_\mu^\nu} \frac{1}{\|kA\|}$ pour tous les polygônes P_μ^ν possibles. On a :

$$\sigma(j) < b_\nu M_\nu^\nu + b_{\nu-1} M_{\nu-1}^\nu + \dots + b_0 M_0^\nu \tag{21}$$

D'après la propriété I, le terme le plus grand de $\sum_{P_\mu^\nu} \frac{1}{\|kA\|}$ est inférieur à $\frac{1}{\delta_\nu}$. Il lui correspond un sommet S_0 du polygône P_μ^ν. Si on appelle $S_1, S_2 \dots$ et $S_1', S_2' \dots$ les sommets de P_μ^ν consécutifs à S_0 dans un sens et dans l'autre, et Ω le point d'affixe 1, on a :

$$\left.\begin{array}{l}\widehat{\Omega S_1} > \widehat{\Omega S_0} \\ \widehat{\Omega S_1'} > \widehat{\Omega S_0}\end{array}\right\} \longrightarrow \begin{array}{l}\widehat{\Omega S_1} > \frac{\widehat{S_0 S_1}}{2} \geqslant 2\pi \frac{\delta_{\mu-1}}{2} \\ \widehat{\Omega S_1'} > \frac{\widehat{S_0 S_1'}}{2} \geqslant 2\pi \frac{\delta_{\mu-1}}{2}\end{array} \quad \text{d'après II}$$



Les termes de $\sum\limits_{P_\mu^\nu}$ correspondants à S_1 et S'_1 peuvent donc être majorés par $\dfrac{1}{\dfrac{\delta_{\mu-1}}{2}}$

Toujours d'après II, les termes de $\sum\limits_{P_\mu^\nu}$ correspondants à S_2 et S'_2 peuvent être majorés par $\dfrac{1}{\frac{1}{2}\delta_{\mu-1} + \delta_{\mu-1}}$ etc.

Finalement :

$$M_\mu^\nu \leqslant \frac{1}{\delta_\nu} + \frac{4}{\delta_{\mu-1}} \left[1 + \frac{1}{3} + \ldots + \frac{1}{Q_\mu - \varepsilon} \right] \qquad (22)$$

$\varepsilon = 2$ si Q_μ est impair

$= 1$ si Q_μ est pair.

Or d'après (17) :

$$\frac{1}{\delta_{\mu-1}} < \frac{1}{\dfrac{1}{Q_\mu} - \dfrac{Q_{\mu-1}}{Q_\mu Q_{\mu+1}}} = \frac{Q_\mu Q_{\mu+1}}{Q_\mu a_{\mu+1}} = Q_\mu + \frac{Q_{\mu-1}}{a_{\mu+1}} < 2\, Q_\mu \qquad (23)$$

et d'autre part :

$$1 + \frac{1}{3} + \ldots + \frac{1}{Q_\mu - \varepsilon} < \frac{1}{2} \left[\log Q_\mu + C + \log 2 + \frac{1}{6 Q_\mu^2} \right] \qquad (24)$$

$$< \frac{1}{2} (\log Q_\mu + G)$$

$G < 1{,}3121$ si $Q_\mu > 1$

(21) devient alors :

$$\sigma(j) < 2\, Q_{\nu+1} \sum_{\mu=1}^{\nu} b_\mu + 4 \sum_{\mu=1}^{\nu} b_\mu Q_\mu (\log Q_\mu + G) \qquad (25)$$

En tenant compte de (19) et (20) :

$$\boxed{\sigma(j) < 2\,Q_{\nu+1} \sum_{\mu=1}^{\nu+1} a_\mu + 4\,j\,(\log j + G)} \qquad (26)$$

III - EVALUATION DE L'ERREUR $|I_N|$

a) A est un nombre irrationnel quadratique

A est une racine irrationnelle d'une équation du second degré à coefficients entiers. On montre que son développement en fraction continue est périodique, c'est-à-dire que la suite a_0, a_1, a_2,... est périodique à partir d'un certain rang. (Théorème de Lagrange : [3] Théorème 177). Il existe donc un nombre B tel que :

$$a_i \leqslant B$$

D'après (14) on a :

$$Q_{\nu+1} < (B+1)\,Q_\nu < (B+1)\,j \qquad (27)$$

et

$$Q_\nu > 2\,Q_{\nu-2} > 2^2 Q_{\nu-4} > \dots > 2^{\frac{\nu}{2}}\,Q_0 \quad \text{ou} \quad 2^{\frac{\nu-1}{2}}\,Q_1$$

c'est-à-dire dans tous les cas :

$$\frac{\nu - 1}{2}\log 2 < \log Q_\nu < \log j \qquad (28)$$

donc :

$$\sum_{\mu=1}^{\nu+1} a_\mu \leqslant B(\nu+1) < \left[\frac{2\log j}{\log 2} + 2\right] B \qquad (29)$$

(26) donne alors :

$$\sigma(j) < 4\,j\log j \left[\frac{B+1}{\log 2} + 1\right] + 4\,j\,[B+1+G] \qquad (30)$$

$$\sigma(j) = O(j\log j) \qquad (31)$$

Or d'après (4)

$$\gamma_j - \gamma_{j+1} < \frac{M(1+\lambda)}{j^{2+\lambda}} = O\left(\frac{1}{j^{2+\lambda}}\right) \qquad (32)$$

La série S intervenant dans (10) est donc convergente et l'on obtient bien :

$$I_N < \frac{K}{N} \tag{33}$$

où K est un nombre que l'on peut calculer facilement en fonction de B, M et λ d'après (30) et (32).

On voit que la périodicité du développement de A n'intervient pas directement : le résultat est le même si A est un nombre dont les quotients incomplets sont bornés par B.

b) A est un nombre irrationnel algébrique quelconque.

Nous démontrerons simplement que la série S est convergente sans chercher à obtenir une majoration de sa somme.

D'après le théorème de Roth ([2] chapitre VI) l'inégalité

$$| QA - P | > \frac{1}{Q^{1+\delta}} \qquad \delta > 0 \tag{34}$$

est valable sauf pour un nombre fini de valeurs de P et Q. D'autre part d'après (16) et (17), il existe un nombre P_n tel que

$$| Q_n A - P_n | < \frac{1}{Q_{n+1}} \tag{35}$$

On en déduit que

$$Q_{n+1} < Q_n^{1+\delta} \tag{36}$$

sauf peut-être pour un nombre fini de valeurs de n.

De (14) et (36) :

$$a_\mu < Q_{\mu-1}^{\delta}$$

$$\sum_{\mu=1}^{\nu+1} a_\mu < \sum_{\mu=1}^{\nu+1} Q_{\mu-1}^{\delta} < 2 [Q_\nu^{\delta} + Q_{\nu-2}^{\delta} + \dots]$$

$$< 2 Q_\nu^{\delta} \left[1 + \frac{1}{2^\delta} + \frac{1}{2^{2\delta}} + \dots \right] < \frac{2^{\delta+1}}{2^\delta - 1} Q_\nu^{\delta}$$

Donc :

$$\sum_{\mu=1}^{\nu+1} a_\mu = O(j^\delta) \tag{37}$$

et par conséquent :

$$\sigma(j) = O(j^{1+2\delta}) \tag{38}$$

Comme on peut choisir δ inférieur à $\frac{\lambda}{2}$, on voit que la série S est convergente et que I_N est bien de l'ordre de $\frac{1}{N}$ lorsque A est un nombre irrationnel algébrique dès que λ est positif.

IV - EXEMPLES NUMERIQUES -

Nous allons étudier plus particulièrement le comportement de I_N en fonction de N pour deux fonctions f(x):

$$f_1(x) = \begin{cases} x & \text{si } 0 \leqslant x \leqslant \frac{1}{2} \\ 1 - x & \text{si } \frac{1}{2} \leqslant x \leqslant 1 \end{cases}$$

$$f_2(x) = \frac{a(1 - a^2)}{1 + a^2 + 2a \cos 2\pi x} \qquad \text{avec } a = \frac{1}{4}$$

et pour trois nombres A :

$$A_1 = \sqrt{2} \; ; \qquad A_2 = \pi \; ; \qquad A_3 = \frac{22}{7}$$

$f_1(x)$ est une fonction continue mais dont la dérivée a une discontinuité de première espèce. Sa série de Fourier converge donc assez lentement :

$$f_1(x) = \frac{1}{4} - \frac{1}{\pi^2} \sum_{h=-\infty}^{+\infty} \frac{1}{(2h + 1)^2} e^{2i\pi(2h+1)x} \tag{39}$$

$f_2(x)$ est une fonction indéfiniment dérivable dont la série de Fourier converge très rapidement :

$$f_2(x) = \frac{1}{4} - \frac{1}{4} \sum_{\substack{k=-\infty \\ k \neq 0}}^{+\infty} \frac{(-1)^k}{4^k} e^{2i\pi kx} \tag{40}$$

Les planches 1 et 2 donnent l'allure des courbes représentant I_N en fonction de N jusqu'à N = 250.

Sur les planches 3 à 8 ont été représentées en coordonnées logarithmiques :

1) Les valeurs de I_N en fonction de N pour chaque valeur de N inférieure à 30. Lorsque cela a été possible sans embrouiller la figure, les points ont été joints.

2) Les valeurs de I_N pour certaines valeurs de N supérieures

à 30, en principe les valeurs correspondantes aux points d'abscisse exacte (ou un point voisin lorsque I_N sort de la figure).

3) Les majorations obtenues dans les paragraphes suivants :

4) Sur les planches 3 et 5, les majorations obtenues par M. Bass dans la référence [1].

a) Etude du cas $A_1 = \sqrt{2}$.

A_1 est un nombre algébrique quadratique dont le développement en fraction continue est très régulier :

$$\hat{A} = 1 \quad \text{et} \quad a_1 = a_2 = \dots = 2$$

D'après les résultats obtenus au §III-a, les séries S et $\mathfrak{S}$ sont convergentes. Donc si on appelle S^* la plus petite des bornes supérieures de $N|I_N|$:

$$S^* = \lim\sup N|I_N|,$$

on est sûr que S^* existe. A l'aide de (10) et de (30) qui donne :

$$\sigma(j) < 21,3\, j \log j + 17,3\, j \,, \tag{41}$$

on peut obtenir un premier majorant M_0 de S^*.

La majoration (41) étant assez grossière, on peut avoir un meilleur majorant M_1 en calculant directement la somme de la série $\mathfrak{S}$, ne se servant des transformations (9) et (10) et de (41) que pour évaluer le reste de la série $\mathfrak{S}$.

Lorsque la fonction périodique de période 1 qui coincide avec f(x) dans l'intervalle (0, 1) est paire, (ce qui est le cas pour f_1 et f_2), on peut encore améliorer facilement la majoration. On décompose (5) en deux sommes :

$$I_N = \frac{1}{N}\left[\sum_{k\neq 0} \frac{c_k}{e^{2i\pi kA} - 1}\, e^{2i\pi kAN} - \sum_{k\neq 0} \frac{c_k}{e^{2i\pi kA} - 1}\right] \tag{42}$$

La première somme est majorée par $\mathfrak{S}$. La seconde somme est égale à :

$$-\sum_{k=1}^{\infty} c_k = -\frac{f(0) - c_0}{2} \begin{cases} = 0,125 & \text{pour } f_1 \\ = 0,05 & \text{pour } f_2 \end{cases}$$

Un troisième majorant de S^* est donc

$$M_2 \doteq \mathcal{S} + \left| \sum_{k=1}^{\infty} c_k \right|$$

Pour la fonction f_2, le calcul de $\mathcal{S}$ est facile : il suffit de calculer les six premiers termes et la majoration du reste donne :

$$0,103 < \mathcal{S} < 0,104$$

Donc $$0,206 < M_1 < 0,208$$

et $$0,153 < M_2 < 0,154$$

Pour la fonction f_1, la série $\mathcal{S}$ converge très lentement ; le calcul a été fait pour les treize premiers termes :

$$0,154 < \mathcal{S}_{13} < 0,155$$

Le terme d'erreur est encore très grand :

$$\mathcal{R}_{13} < 0,88$$

On obtient donc :

$$N\,|I_N| < 2,07$$

ou un peu mieux $N\,|I_N| < 1,16.$ *(Voir Note à la fin)*

Sur la planche 3 ont été tracées les hyperboles H_m, majorations obtenues dans [1]. Les courbes sont régulièrement disposées et correspondent à une majoration de la forme

$$|I_N| < \frac{K}{\sqrt{N}}$$

b) Etude du cas $A_2 = \pi$

A est un nombre transcendant dont on ne connaît que les premiers termes du développement en fraction continue.

$$\hat{A} = 3, \quad a_1 = 7, \quad a_2 = 15, \quad a_3 = 1, \quad a_4 = 292, \ldots$$

Ce développement est assez irrégulier et les quotients incomplets sont relativement grands.

On ne sait pas si les séries S et $\mathcal{S}$ convergent. On peut quand

même utiliser les résultats du §II pour trouver une majoration de I_N. On sépare I_N en plusieurs morceaux :

$$I_N = \frac{1}{N} \sum_{\substack{k=-Q_m \\ k\neq 0}}^{k=+Q_m} c_k \sum_{n=0}^{N-1} e^{2i\pi k A n} + \sum_{|k|>Q_m} c_k \frac{1}{N} \sum_{n=0}^{N-1} e^{2i\pi k A n} \qquad (43)$$

Le second terme α_1 peut être majoré par :

$$\sum_{|k|>Q_m} |c_k| = \varepsilon_m$$

qu'on peut rendre petit en choisissant Q_m assez grand. Le premier terme peut être encore décomposé puis majoré comme en (7) et (8). On obtient :

$$|I_N| < \varepsilon_m + \frac{2}{N}\left[\sum_{k=1}^{k_0} \frac{|c_k|}{|\sin \pi k A|} + \sum_{j=k_0+1}^{Q_1-1} [\gamma(j) - \gamma(j+1)]\sigma(j) + \sum_{Q_1}^{Q_2-1} + \ldots \right.$$

$$\left. \ldots + \sum_{Q_{m-1}}^{Q_m-1} [\gamma(j) - \gamma(j+1)]\sigma(j)\right]. \qquad (44)$$

$$= \varepsilon_m + \frac{2}{N}[\alpha_2 + \alpha_3 + \ldots \alpha_{m+2}]$$

Le premier terme α_2 se calcule directement et les autres termes se majorent en utilisant la formule (25) correspondant à chaque somme.

Pour le calcul numérique, on a pris $k_0 = Q_1 = 7$ et $Q_m = Q_2 = 106$. On obtient :

pour f_1 = $|I_N| < \dfrac{3,17}{N} + 0,000965$ *(Voir Note à la fin)*

pour f_2 = $|I_N| < \dfrac{0,347}{N} + 10^{-60}$

c) <u>Etude du Cas A_3</u> $= \dfrac{22}{7}$

A_3 est un nombre rationnel qui représente la première "bonne" approximation rationnelle de π.

Lorsque A est rationnel, la moyenne J_N ne tend en général pas vers l'intégrale J. I_N tend vers un nombre l différent de 0. Ici :

$$l = \frac{1}{7} \sum_{n=0}^{6} f\left(\underset{\smile}{\frac{22}{7}\, n}\right) - J \qquad (45)$$

ce qui donne :

$$\text{pour } f_1 = l_1 = \frac{12}{49} - \frac{1}{4} = -0,0051\ldots$$

$$\text{pour } f_2 = l_2 = 0,2499694.. - 0,25 = -0,0000305\ldots$$

Le nombre l a aussi pour valeur :

$$l = \sum_{h=1}^{\infty} c_{7h} \tag{46}$$

La manière dont se fait la convergence de I_N vers l est facile à étudier exactement ; par exemple pour $f(x) = f_1(x)$:

$$\begin{aligned}
I_N &= l_1 = -0,0051\ldots && \text{si } N \equiv 0 \text{ modulo } 7\\
&= l_1 - \frac{12}{49}\frac{1}{N} && \text{si } N \equiv 1 \quad ''\\
&= l_1 - \frac{17}{49}\frac{1}{N} && \text{si } N \equiv 2 \quad ''\\
&= l_1 - \frac{15}{49}\frac{1}{N} && \text{si } N \equiv 3 \quad ''\\
&= l_1 - \frac{6}{49}\frac{1}{N} && \text{si } N \equiv 4 \quad ''\\
&= l_1 + \frac{3}{49}\frac{1}{N} && \text{si } N \equiv 5 \quad ''\\
&= l_1 + \frac{5}{49}\frac{1}{N} && \text{si } N \equiv 6 \quad ''
\end{aligned}$$

Ces courbes ont été tracées sur les planches 7 et 8. Sur la figure 8 ce sont presque des droites car l_2 est très petit.

Il est intéressant d'appliquer au cas $A = A_3$ la méthode utilisée en b). On prend $\dot{Q}_m = 7$ et $k_0 = 6$.

$$|I_N| < \frac{2\,\alpha_2}{N} + |\alpha_1|$$

$$|\alpha_1| < \varepsilon_m = \sum_{k=7}^{\infty} |c_k|$$

qui est bien supérieur à l.

$$2\,\alpha_2 = \begin{cases} 0,500 & \text{pour } f_1 \\ 0,330 & \text{pour } f_2 \end{cases}$$

Donc :

$$\text{pour } f_1 = |I_N| < \frac{0,501}{N} + 0,0169$$

$$\text{pour } f_2 = |I_N| < \frac{0,331}{N} + 0,000041$$

d) Remarques.

1) La régularité de la fonction f(x) ne paraît pas avoir beaucoup d'influence sur la forme des courbes I_N en fonction de N, ni sur la valeur de

$$S^* = \lim\sup N|I_N|$$

Elle paraît influer davantage sur la manière dont cette valeur est approchée.

2) La formule (42) montre que NI_N est une fonction périodique de la variable continue N. Une étude plus approfondie montrerait que la fonction continue ayant pour valeur NI_N lorsque N et entier et variant linéairement entre les points d'abscisses entières est une fonction presque périodique. Un système de presque-périodes est l'ensemble des dénominateurs Q_m. On trouve ainsi l'explication de l'allure des courbes I_N en fonction de N (planches 1 et 2).

- Allure irrégulière pour $A_1 = \sqrt{2}$.

Les périodicités sont peu apparentes car les presque-périodes sont très rapprochées :

$$Q_1 = 2, \; Q_2 = 5, \; Q_3 = 12, \; Q_4 = 29, ..$$

- Périodicités très apparentes pour A_2 et A_3.

car les presque-périodes sont bien séparées. Pour A_3 il y a une période exactement égale à 7 et pour A_2 la presque-période de longueur $7 = Q_1$ se superpose à la presque-période de longueur $106 = Q_2$ cette dernière visible surtout pour $f = f_1$ (planche 1).

REFERENCES

[1] J. BASS - Nombres aléatoires. Suites arithmétiques. Méthode de Monte-Carlo. Publications de l'Institut de Statistique de l'Université de Paris. Vol. IX - Fas. 3, p. 289-325.

[2] J.W.S. CASSELS - An introduction to diophantine approximations. Cambridge University Press. (1957).

[3] G.H. HARDY and E.M. WRIGHT - An introduction to the theory of numbers. Oxford Clarendon Press. (1938).

[4] H. WEYL - Über die Gleichverteilung von Zahlen modulo Eins (Math. Annalen, 77, 1916, p. 313-352).

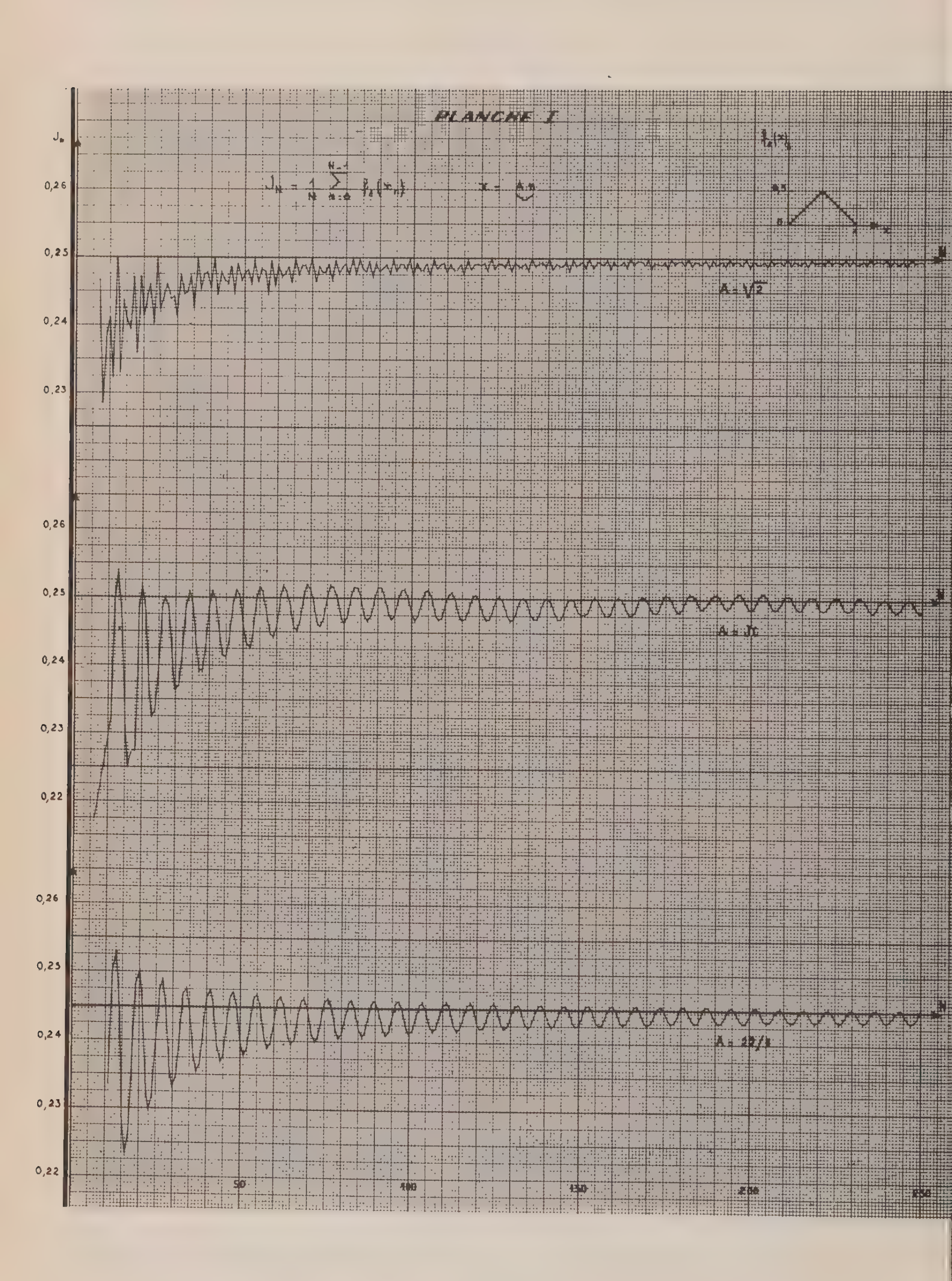
Note : Résultats du calcul numérique.

$f(x) = f_1(x) \quad A_1 = \sqrt{2}$

Le maximum de $N|I_N|$ pour $N < 5.10^5$ est 0,25758..., valeur atteinte 216 fois, la première fois pour N = 5916, la dernière pour N = 491 448.

$f(x) = f_1(x) \quad A_2 = \pi$

Le maximum de $N|I_N|$ pour $N < 6,5.10^5$ est 0,48552..., valeur atteinte 30 fois, la première fois pour N = 8272, la dernière pour N = 638 574.


PLANCHE I
J_N
$J_N = \frac{1}{N} \sum_{n=0}^{N-1} f_2(x_n)$
0,26
0,25
0,24
0,23
0,22
A = $\sqrt{2}$
A = π
50
100
150
200
250

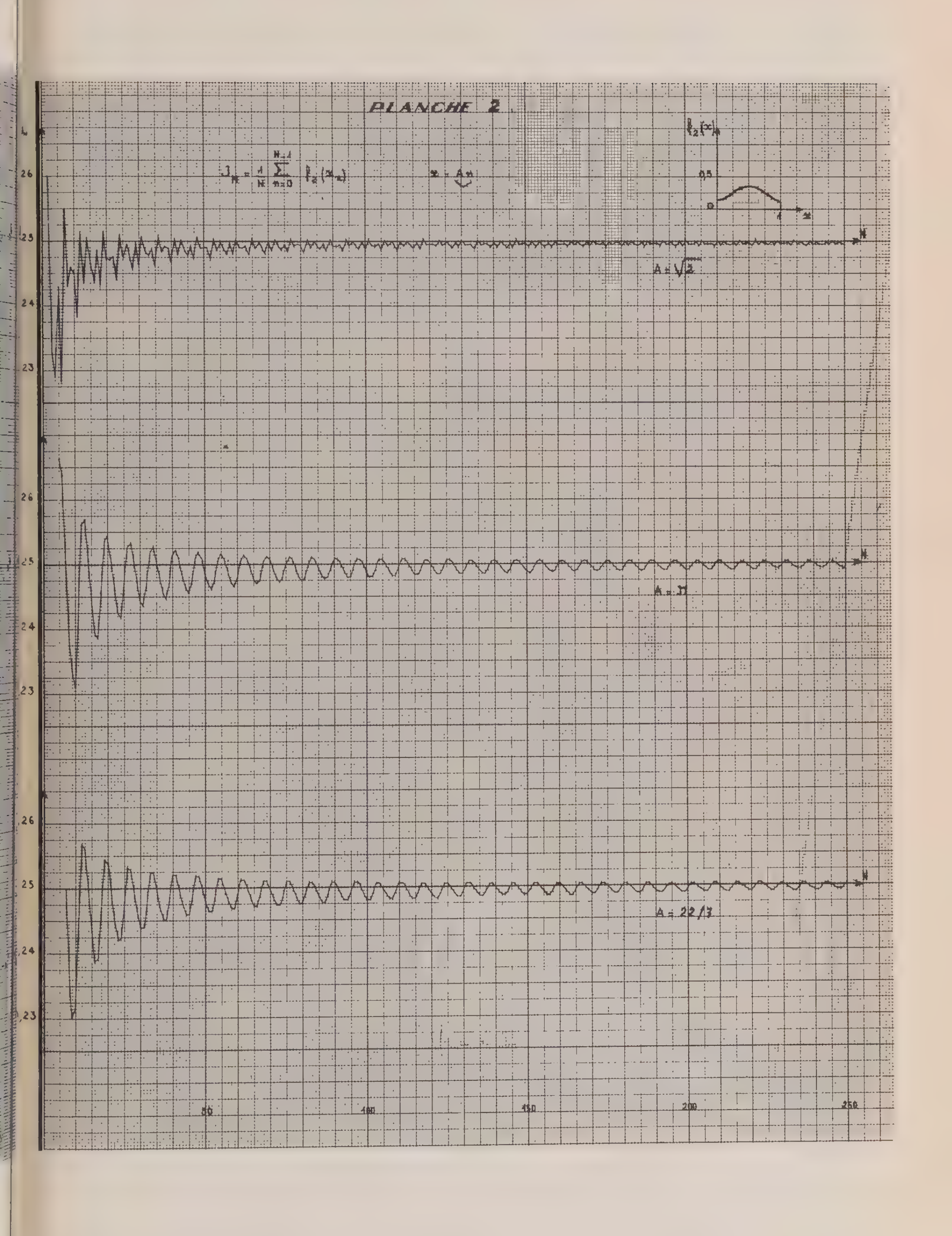

PLANCHE 2
$J_N = \frac{1}{N} \sum_{n=0}^{N-1} f_2(x_n)$
$x_n = \widetilde{An}$
$f_2(x)$
0,5
0
1
x
N
A = √2
A = π
A = 22/7
,26
,25
,24
,23
50
100
150
200
250

PLANCHE III

PLANCHE 3

$A = \sqrt{2}$

majoration 1

majoration 2

I_N

N

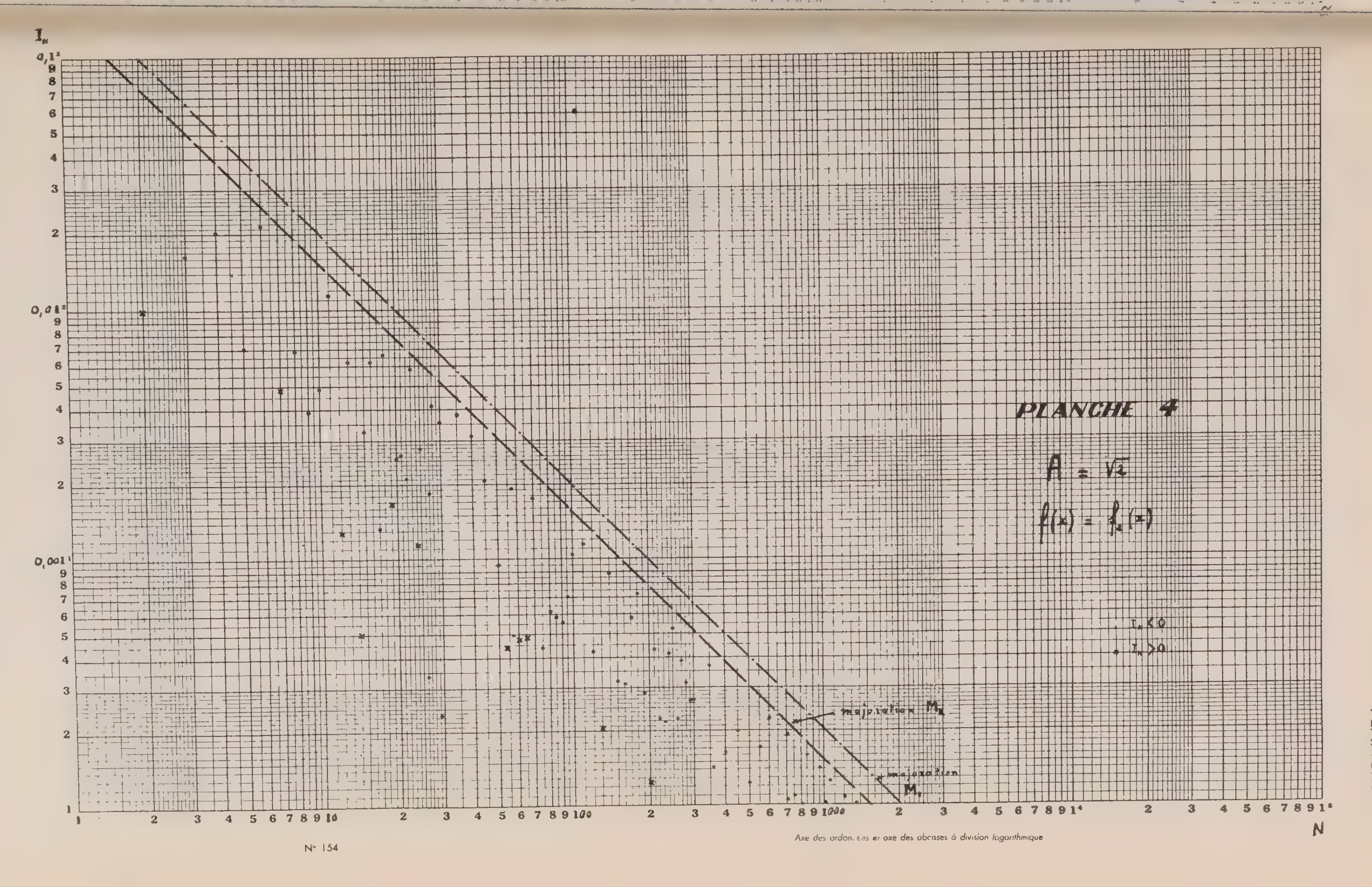
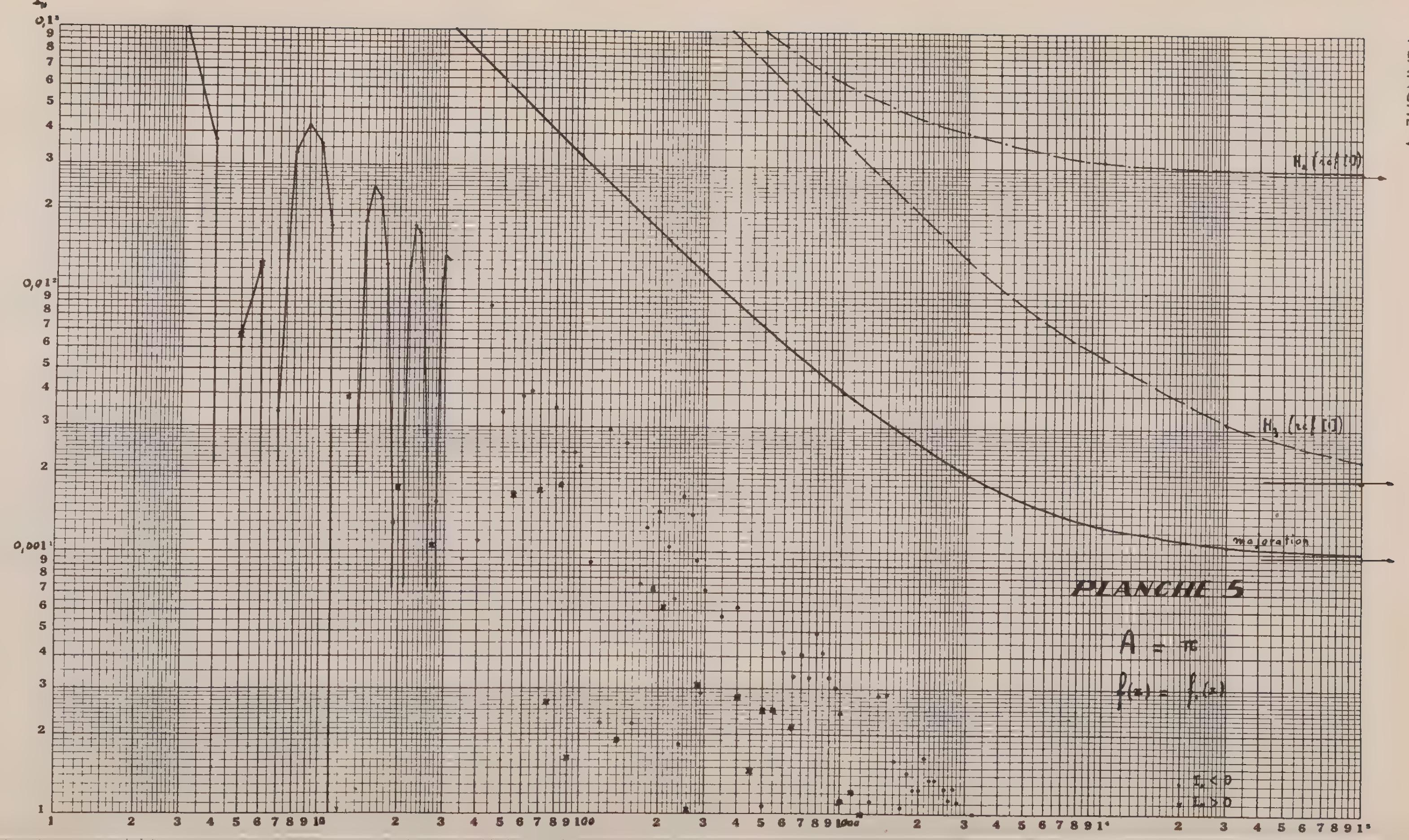

PLANCHE 4
$f(x) = f_1(x)$
$I_N < 0$
$I_N > 0$
majoration M_2
majoration M_1
I_N
N
Axe des ordonnées et axe des abcisses à division logarithmique
N° 154

PLANCHE V



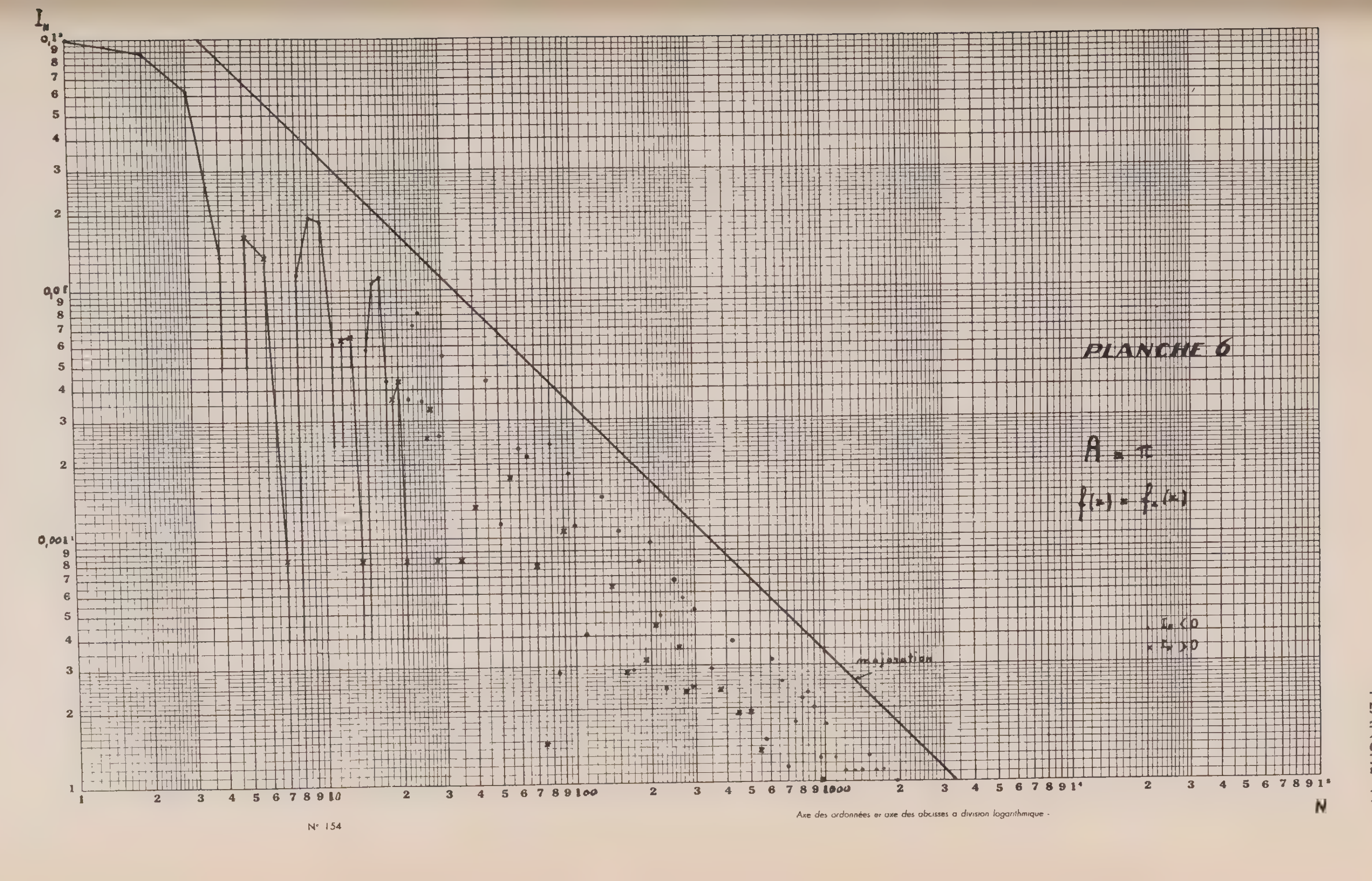
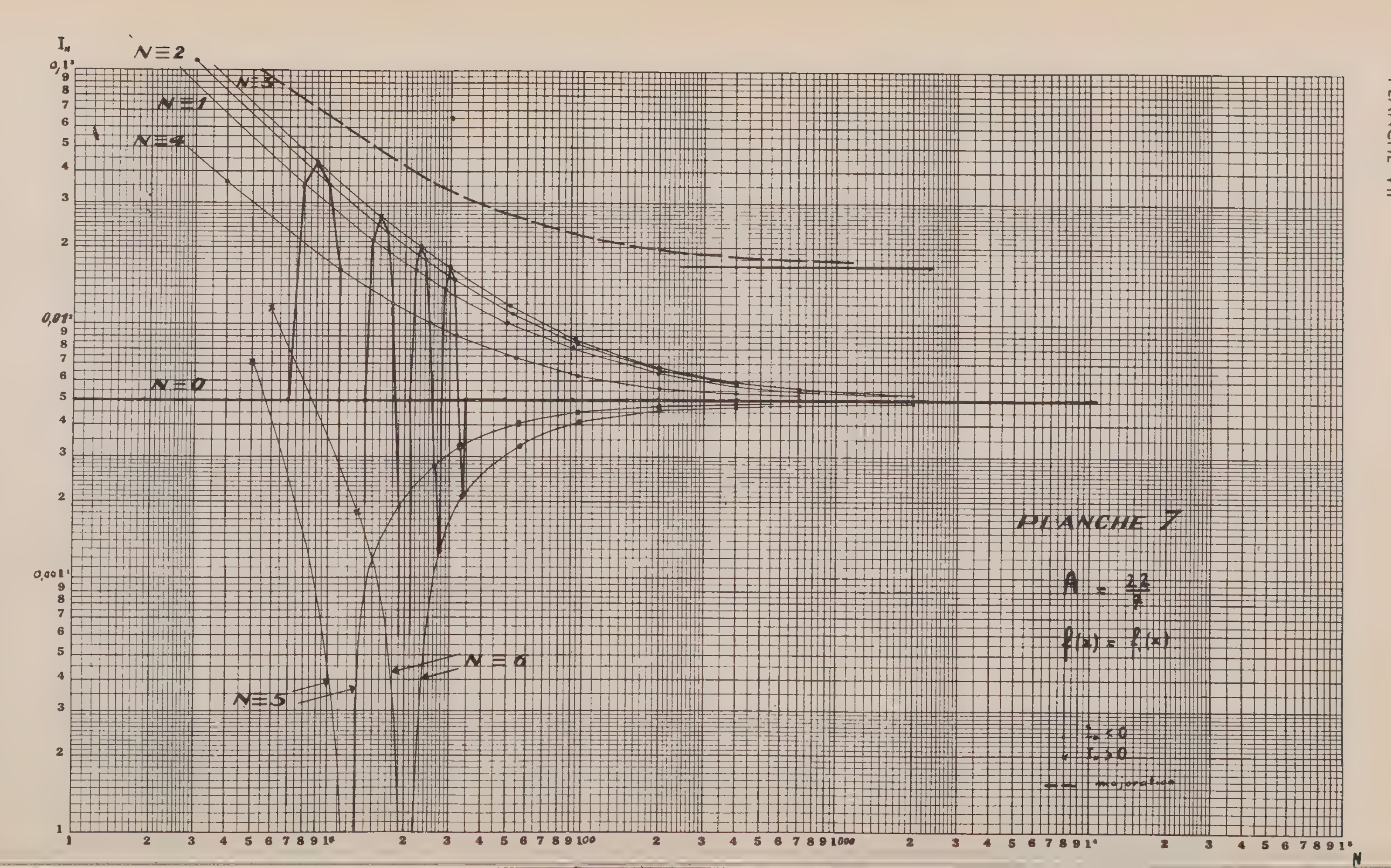
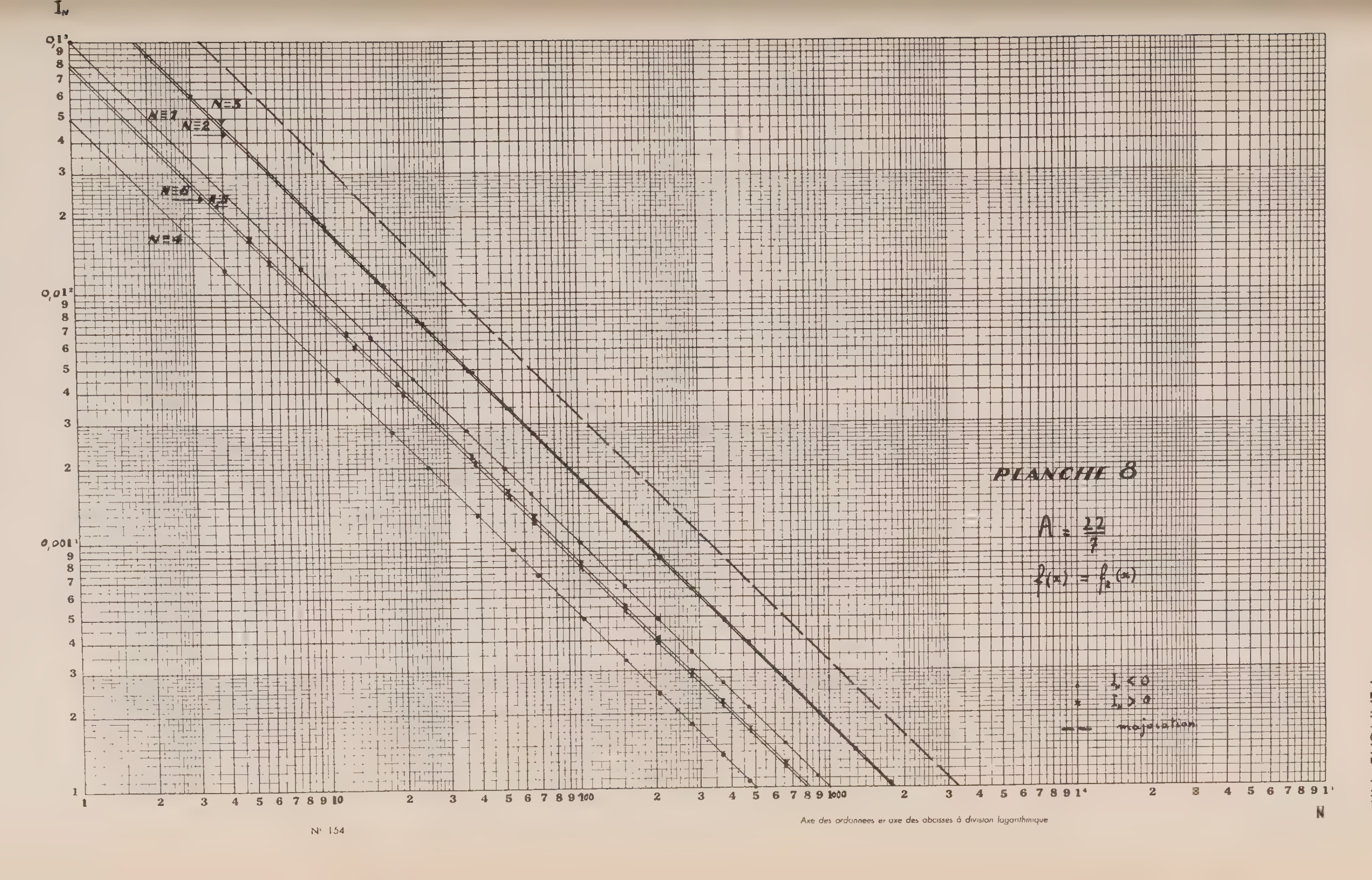

PLANCHE 6
A = π
f(x) = f₁(x)
• I_N < 0
x I_N > 0
majoration
I_N
N
Axe des ordonnées et axe des abscisses à division logarithmique -
N° 154

PLANCHE VII




PLANCHE 8
A = 22/7
f(x) = f2(x)
I_N < 0
I_N > 0
majoration
Axe des ordonnées et axe des abcisses à division logarithmique
N° 154
N
I_N
N=3
N=2
N=1
N=6
N=4

EXTENSION DU CONCEPT DE CONVERGENCE EN PROBABILITÉ AUX VARIABLES ORDINAIRES

Vittorio CASTELLANO

RESUME

Ayant défini l'ordre infini d'un point d'accumulation d'une suite et les points principaux d'accumulation ; l'auteur définit la convergence presque exclusive, la convergence d'une suite de nombres réels à un point-limite principal et il étend cette forme de convergence au cas d'une suite de distributions de nombres réels vers une distribution limite principale.

De même que la convergence ordinaire, ou exclusive à une seule limite, ainsi la convergence à une distribution limite principale peut être convergence "partout" ou "presque partout" selon qu'elle se vérifie pour tous les points de la distribution sans exception, ou bien pour tous les points excepté un ensemble de dimension zéro.

De ce mode très général de convergence vers un point-limite principal est la tendance vers une constante de la variable de Bernoulli, lorsque le nombre d'épreuves croît indéfiniment, qui aussi se réduirait sur le terrain de l'Analyse classique.

La "convergence en probabilité" de Fréchet, sous la forme explicite la plus avantageuse, est de même un cas particulier de la convergence "presque exclusive", qui étend aux variables de l'Analyse classique cette particulière "convergence vers une limite" des variables aléatoires, très efficacement étudiée, par Cantelli et d'autres après lui, avec le nom de convergence vers une limite au sens du calcul des probabilités.

I - L'OBJET DE CETTE ETUDE -

La variable de Bernoulli, que nous indiquons par $\mathfrak{F}_n$ = fréquence des résultats favorables en n épreuves faites à probabilité constante P, nous a donné l'idée d'une convergence différente de celle de l'analyse mathématique, que l'on a nommée convergence "au sens du calcul des probabilités", ou "en probabilité", ou "stocastique", selon les différents auteurs.

Cette convergence est représentée par l'inégalité

$$\text{Prob}\,\{\, |\, \mathfrak{F}_n - P\, | < \varepsilon \,\} > 1 - \eta \qquad (1)$$

valable dès que n est assez grand pour tout couple ε, η arbitrairement petits et positifs, mais si l'on calcule, pour chaque valeur déterminée de n, les résultats possibles $\mathfrak{F}_n$ satisfaisant l'inégalité

$$|\, \mathfrak{F}_n - P\, | < \varepsilon$$

on peut écrire aussi, au lieu de la (1) :

$$\text{Freq}\,\{\, |\, \mathfrak{F}_n - P\, | < \varepsilon \,\} > 1 - \eta \qquad (2)$$

Dans notre étude nous nous en tiendrons, généralement, à cette forme d'écriture qui, à la probabilité d'un événement, substitue sa fréquence dans l'ensemble des cas possibles, ou "univers des épreuves" [1].

Le but de notre étude est de démontrer que la convergence à une limite de la variable de Bernoulli appartient à un type général de convergence "presque exclusive", pouvant être définie aussi pour la suite des nombres réels ; elle peut être étudiée au moyen de l'Analyse mathématique, indépendamment du concept de variable aléatoire.

II - CONVERGENCE "PRESQUE EXCLUSIVE" D'UNE SUITE DE NOMBRES REELS -

Nous tâchons d'abord de transcrire (2) au cas de suites de nombres réels. Comme le $\mathfrak{F}_n$ est une variable, il faut transformer la suite des nombres réels :

$$a_1,\ a_2, \ldots\ldots, a_n, \ldots\ldots \qquad (3)$$

en une suite de variables réelles

$$A_1,\ A_2, \ldots\ldots, A_n, \ldots\ldots \qquad (3')$$

et nous pouvons faire cela en définissant

$$A_i = A_i(a_i,\ a_{i+1}, \ldots\ldots)$$

----

(1) V. Castellano : "Sull'universo delle prove come fondamento della teoria dei campioni", 31e Session de l'Institut International de Statistique, Bruxelles, 1958.

comme l'ensemble de tous les termes de (3) de rang supérieur à i.

La transcription de (2) par (3) :

$$\text{Freq}\ \{ |A_n - a| < \varepsilon \} > 1 - \eta \qquad (2')$$

signifie que, posant $K(n\,n')$ = nombre des termes suivant de (3) :

$$a_{n+1},\ a_{n+2}, \ldots\ldots, a_{n'}$$

n'étant pas dans l'intervalle $a \mp \varepsilon$, pour un certain n assez grand, il y a l'inégalité

$$K(n, n') < (n' - n)\,\eta,$$

et, par conséquent, que

$$\lim_{n' \to \infty} \frac{K(n, n')}{n' - n} = 0$$

Nous exprimons ce fait en disant que l'ensemble de ces $K(n, n')$ points est infini d'ordre inférieur relativement à l'ordre infini de l'ensemble total que nous supposons = 1.

Comme $K(0, n') < n + K(n, n')$, on a aussi $\lim_{n' \to \infty} \frac{K(0, n')}{n'} = 0$ et par conséquent tous les termes de (3), en dehors de l'intervalle $a \pm \varepsilon$, constituent un sous-ensemble de (3) d'ordre infini < 1 en comparaison de l'ensemble (3).

Nous indiquons plus simplement par $K(n)$ le nombre des termes de cet ensemble parmi les premiers n termes de la (3).

Etant donné cela, nous disons que la suite (3) converge vers la limite "principale" a et nous écrivons :

$$\lim_{n \to \infty}{}_{p.}\ a_n = a \qquad (4)$$

s'il arrive que :

$$\text{Freq}\ \{ |a_n - a| < \varepsilon \}\ \ 1 - \eta \qquad (5)$$

c'est-à-dire si toute la suite appartient à l'intervalle $a \pm \varepsilon$, sauf un sous-ensemble infini d'ordre < 1.

Si de $n > n_0$ assez grand

$$K(n, n') = 0$$

pour tout n' , on a

$$\text{Freq} \{ | a_n - a | < \varepsilon \} = 1 \qquad (6)$$

en toutes les

$$A_{n_0 , n'} = A_{n'} - A_{n_0}$$

et a_n tend vers a dans le sens ordinaire de l'Analyse.

D'après les définitions que nous avons données, (5) définit un type de convergence qui diffère de la convergence considérée ordinairement par l'Analyse, à cause de l'existence de plusieurs points d'accumulation au lieu d'un seul point. Mais un seul de ces points a une valeur numérique du même ordre infini que la suite totale, et nous l'appellerons "point d'accumulation principale".

Si, par exemple, la suite des valeurs réciproques des nombres naturels est :

$$1, \ 1/2, \ 1/3, \dots, \ 1/n \qquad (7)$$

celle-ci converge à la limite zéro.

La suite

$$1, \ 2, \ 1/3, \ 2^2, \ 1/5, \ 1/6, \ 1/7, \ 2^3, \ 1/9, \dots \qquad (8)$$

a deux points d'accumulation : zéro et ∞.

Dans les premiers n termes d'elle, il y a dehors de l'intervalle (0, 1) un nombre $K(n)$ de termes égaux au maximum entier contenu en $\log_2 n$ et comme

$$\lim_{n \to \infty} \frac{\log_2 n}{n} = 0 \qquad (9)$$

zéro est le point d'accumulation principale et (8) a zéro comme limite principale.

Les considérations précédentes sont valables pour tout ensemble d'un nombre fini de valeurs et, en général, pour toute variable

dont la distribution des fréquences est connue. Si nous supposons, en général, comme définition de convergence d'une variable X_n à une limite finie a, la (2).

$$\text{Freq}\,\{\,|\,X_n - a\,| < \varepsilon\,\} > 1 - \eta \tag{10}$$

il s'ensuit immédiatement que : si une suite de variables réelles converge vers une limite déterminée et finie, cette limite est unique ou elle est principale, selon que

$$\lim_{n\to\infty} w_n \begin{smallmatrix}=\\ \neq\end{smallmatrix} 0 \tag{11}$$

où w_n est le champ de variation de X_n.

III - OBSERVATIONS PRELIMINAIRES SUR LA REPRESENTATION D'UNE VARIABLE -

Avant d'appliquer les définitions, que nous venons de donner, à une suite de variables réelles, tendant vers une variable-limite, nous faisons quelques observations sur les représentations ordinaires des variables réelles.

Les variables réelles ne sont pas des ensembles ordonnés en soi, mais elles peuvent certainement être ordonnées selon l'intensité des valeurs qu'elles acquièrent. Cette distribution des valeurs est représentée, sur le plan carthésien, indifféremment par la courbe de graduation ou par la fonction de répartition, selon que sur l'axe des coordonnées on met les valeurs de la variable ou les fréquences des valeurs ne surpassant pas une certaine limite, seulement si l'on considère, évidemment, les fonctions de répartition en leur propre champ d'existence, au lieu de les considérer dans le domaine le plus vaste possible de $-\infty$ à $+\infty$, comme le font plusieurs auteurs.

Si la variable est donnée sous une forme continue, il faut toujours supposer l'ensemble comme infini, car la masse est distribuée en petits intervalles infinitésimaux dx.

Ayant convenu cela, la constante $X=a$ a comme fonction de répartition

$$F(X) = \begin{cases} \emptyset \\ 0 - 1 \quad \text{pour} \quad x \\ \emptyset \end{cases} \begin{cases} < a \\ = a \\ > a \end{cases} \tag{12}$$

tandis que la "presque constante" zéro représentée par la (7), à la fonction de répartition

$$F(X) = \begin{cases} \emptyset & \\ 0 & \\ 1 & \text{pour} \\ \emptyset & \end{cases} \begin{matrix} x < 0 \\ x = 0 \\ 0 < x \leqslant 1 \\ 1 < x \end{matrix} \qquad (13)$$

où le symbole $\emptyset$ signifie : inexistant.

La suite (8) aussi, dont la fonction de répartition existe en tout le champ $x \geqslant 0$, représente zéro comme une "presque constante", mais sa fonction de répartition est :

$$F(X) = \begin{cases} \emptyset \\ 0 \\ 1 \end{cases} \text{pour} \quad x \begin{cases} < 0 \\ = 0 \\ > 0 \end{cases} \qquad (14)$$

Nous indiquons maintenant par $X(F)$ la variable ordonnée X, c'est-à-dire sa courbe de graduation, et par F(X) sa fonction de répartition. Entre les fonctions correspondantes, que par commodité nous écrivons avec les notations du cas continu, qui inclut, par les conventions ordinaires, le cas discontinu, $x(F)$ et $F(x)$, on a la relation

$$x(F) = F^{-1}(x) \qquad (15)$$

IV - CONVERGENCE "PRESQUE EXCLUSIVE" D'UNE SUITE DE VARIABLES REELLES -

Nous pouvons maintenant appliquer la (5) au cas général de la convergence d'une suite de variables réelles à une variable-limite principale.

Définition :

Nous disons que la suite de variables réelles $X_i (i=1,2,\ldots,n,\ldots)$ a la variable X comme limite presque exclusive ou principale, si, étant donné un $\varepsilon > 0$ pour tous les X_i, sauf un sous-ensemble infini d'ordre < 1, l'inégalité

$$| X_i(F) - X(F) | < \varepsilon \qquad (16)$$

est satisfaite pour chaque valeur de F, sauf tout au plus un ensemble infini de ces valeurs, de dimension nulle.

Si le sous-ensemble de variables X_i, qui fait exception, est fini ou nul, la limite est unique et la convergence est ordinaire.

Soit la convergence ordinaire, soit la convergence principale peuvent avoir lieu "partout" ou "presque partout", selon que (16) est valable pour toutes les valeurs de F sans exception, ou avec l'exception d'un ensemble de dimension nulle.

La convergence presque exclusive partout est, par conséquent, définie par la même inégalité

$$\text{Freq}\ \{\ |X_n(F) - X(F)| < \varepsilon\ \} > 1 - \eta \qquad (17)$$

définissant la convergence des suites de nombres réels, la fréquence $1 - \eta$ étant relative à l'ensemble des n premiers X_i de la suite des X_i, pour $n > n_0$. Dans la convergence presque exclusive, presque partout, la fréquence $1 - \eta$ est relative à l'ensemble des X_i, et pour un certain n assez grand, pour chaque X_i satisfaisant (16), à l'ensemble des valeurs de F. Quand il y a une possibilité d'équivoque, il faut distinguer les différents cas, en mettant l'indication de la variable en question en bas, avant les parenthèses, c'est-à-dire qu'il faut écrire :

$$\text{Freq}_n \qquad \text{ou bien} \quad \text{Freq}_{n.F} \qquad \text{au lieu de} \quad \text{Freq}.$$

La notion Freq_F devrait indiquer la convergence ordinaire presque partout.

Il est facile de construire des exemples de suites de variables montrant ces types de convergence à une limite seule. Si, par exemple, X_n est constitué par l'ensemble des points de l'intervalle $\pm (1 + a_n)$ avec densité

$$y = C_n \sqrt{(1 + a_n)^2 - x^2}$$

la suite des X_n converge partout, à la variable, dont la distribution de la fréquence est

$$y = C \sqrt{1 - x^2}$$

exclusivement ou presque exclusivement, selon que les a_n sont des valeurs de la suite (7) ou de la (8).

D'une manière analogue, la distribution normale :

$$y = \frac{1}{\sigma \sqrt{2}}\, e^{-\frac{(x-M)^2}{2\sigma^2}} \qquad (18)$$

car

$$\lim_{\to 0} y = \begin{matrix} 0 \\ \infty \end{matrix} \quad \text{pour} \quad x \begin{matrix} \neq \\ = \end{matrix} M$$

définit la variable normale tendant vers la presque constante $X = M$, définie dans tout le champ $\mp \infty$, exclusivement ou principalement, selon que l'on donne à σ_1 des valeurs d'une suite convergeant vers zéro comme (7) ou comme (8). Mais si nous considérons la constante M pour chaque σ suffisamment petit, il en résulte :

$$\text{Freq} \{ | X_\sigma - M | < \varepsilon \} > 1 - \eta \tag{19}$$

et par conséquent on peut dire aussi que la variable normale X_σ converge "presque partout" vers la moyenne M, de quelque façon le σ tende vers zéro. Particulièrement la variable de Bernoulli $\mathfrak{F}_n$ converge au sens ordinaire, avec l'augmentation de n, vers la presque constante P, définie dans l'intervalle (0, 1).

Des considérations faites il s'ensuit que si nous mettons

$$\omega (X_n, X) = \max | X_n(F) - X(F) | \tag{20}$$

la condition nécessaire et suffisante, pour que X_n converge au sens ordinaire vers X, est que :

$$\lim_{n \to \infty} \omega(X_n, X) = 0 \tag{21}$$

et que, par conséquent, l'inégalité

$$\lim_{n \to \infty} \omega(X_n, X) > 0 \tag{22}$$

caractérise, avec (17), la convergence partout vers une limite principale, et la convergence presque partout vers une limite exclusive ou principale.

Nous observons que la convergence ordinaire, "presque partout" est une caractéristique des variables aléatoires théoriques, sujet du calcul des probabilités, et que les variables empiriques correspondantes convergent - par la loi empirique du cas - vers la valeur limite théorique, mais seulement "principalement". Par exemple, la distribution d'un nombre suffisamment élevé K de déterminations empiriques indépendantes de la variable de Bernoulli, faites chacune sur n épreuves, a pour $n \to \infty$, la "presque constante" P" comme limite principale, car pour chaque valeur de n, ces K déterminations peuvent même toutes présenter des valeurs bien distantes de P, bien que ce fait soit absolument exceptionnel.

V - CAS PARTICULIER DES VARIABLES DEPENDANTES -

Nous considérons maintenant, d'après les résultats précédents, la signification de l'inégalité

$$\text{Prob}\{ | X_n - X | < \varepsilon \} > 1 - \eta \qquad (23)$$

prise par Fréchet comme la définition de la convergence en probabilité.

Si X_n, et X sont indépendantes, X est nécessairement une constante ou une presque-constante.

Tandis que s'il y a une connexion entre les valeurs des deux variables, alors une partie seule des valeurs de X_n peuvent être associées à un x déterminé, et si l'on remplace aussi Prob par Freq, comme nous avons fait à (1), (23) devrait être écrite plus correctement de la manière suivante :

$$\text{Freq}\{ | X_n(x) - x | < \varepsilon \} > 1 - \eta \qquad (24)$$

pour x variable en tout le champ d'existence de X.

Cela signifie que pour chaque x constant, la variable $X_n(x)$ converge vers la constante certaine x. C'est-à-dire $X_n(x)$ est de la forme

$$X_n = x + y_n(x) \qquad (25)$$

et en définitive

$$X_n = X + Y_n(x) \qquad (25')$$

où $Y_n(x)$ représente un ensemble de fonctions du paramètre x tendant toutes vers zéro quand n tend vers l'infini.

On peut alors l'écrire aussi :

$$\text{Freq}\{ | Y_n(x) | < \varepsilon \} > 1 - \eta \qquad (26)$$

et ainsi que l'équivalente (24), elle rentre comme cas particulier dans la forme plus générale (17) de la convergence presque exclusive, que Fréchet appelle "convergence légale", et que nous pensons qu'elle puisse exprimer le concept essentiel de la convergence en probabilité.

VI - CONCLUSIONS -

A plus de 30 ans de distance de la 1ère édition du "Calcul des Probabilités" par Castelnuovo, où pour la première fois on affirme clairement que le théorème de Bernoulli ne peut pas remplacer la loi empirique du cas dans la légitimation des applications du calcul des probabilités, les auteurs de traités, qui s'adonnent à cette science, continuent à chérir - du moins chez les autres - l'illusion, pas encore déracinée, que le calcul des probabilités étudie les variables aléatoires empiriques. Cela donne au calcul des probabilités l'appât dangereux des choses miraculeuses et impose au statisticiens la mise en lumière des anneaux de passage entre les théories abstraites, la statistique théorique et les applications que les mathématiciens peuvent même négliger.

Si la connexion de la théorie mathématique des probabilités avec les applications est constituée par la loi empirique du cas, l'anneau de passage, pas aperçu toujours, entre le calcul des probabilités abstrait et la statistique théorique, est, selon nous, l'opération aléatoire définissant en bloc et "a priori" la variable aléatoire comme un ensemble de résultats équivalents possibles. La distribution de ceux-ci, comme elle est représentée par la courbe de graduation ou par la fonction de répartition, ou bien par la fonction de densité, définit parfaitement la variable aléatoire, et deux variables aléatoires sont égales, si elles ont la même fonction de répartition. L'opinion contraire de Fréchet est justifiée par la considération de deux variables aléatoires, ayant la même fonction de répartition, qui présentent de différentes déterminations associées, comme, par exemple, il arriverait si elles étaient définies par l'opération de jet d'un seul dé et ensuite par la lecture de deux faces différentes déterminées.

Mais en ce cas, les deux variables sont dépendantes et l'opération aléatoire définissant l'une des deux, Y comme dépendant de l'autre, X, comprend comme composante essentielle l'opération $\vartheta(X)$ qui définit le X, selon le schéma

$$\vartheta(Y) = \vartheta(X).\ \vartheta(x/Y)$$

et l'on ne peut supposer sa fonction de répartition sous la forme de $F(Y)$, mais sous la forme de $F(x/Y)$.

Rien n'empêche de définir égales deux variables Y et X associées et d'écrire :

$$Y(x) = X \quad \text{ou} \quad X(y) = Y$$

quand

$$Y(x) = x$$

ou

$$X(y) = y$$

il serait peut être mieux d'écrire

$$X \bar{=} Y ;$$

mais cela ne signifie point que c'est impossible de définir égales deux variables simples indépendantes et d'écrire :

$$X = Y$$

quand les fonctions de répartition de celles-ci sont égales.

Il est évident que deux variables simples égales ne peuvent pas être coïncidentes et afin qu'elles deviennent associées, il faut introduire de nouvelles conditions qui ne sont pas nécessaires aux variables indépendantes.

L'opinion de V. Mises semble donc être justifiée, c'est-à-dire que les variables aléatoires sont tout à fait définies par leurs fonctions de répartitions, à condition, évidemment, que les variables doubles soient interprétées comme définies par la fonction de répartition double et ainsi de suite.

L'avantage d'une proposition du Calcul des Probabilités, qui soit fondée explicitement sur l'ensemble des résultats possibles de l'opération aléatoire, ou "univers des épreuves" (1), réside, selon nous, dans le fait qu'elle pourvoit la Statistique théorique d'un modèle concret des phénomènes empiriques, sujets à des variations non contrôlables une à une, et encore dans le fait qu'elle pose clairement les limites infranchissables du modèle mathématique, qui ne peut représenter la suite des phénomènes empiriques obéissant aux lois du cas - ou que l'on suppose qu'ils obéissent - mais elle représente seulement les ensembles constituées par ces suites et par leurs parties continues.

Dans notre étude, nous présentons la convergence des suites vers une limite principale comme le modèle mathématique de la convergence vers une limite des collectives empiriques.

----

(1) V. Castellano : "Statistica e Matematica", Discours d'ouverture de XVIIIe Réunion de la Société Italienne de Statistique, Roma, 1958.

Cette convergence, que nous appelons "presque exclusive" en opposition à la convergence "ordinaire" vers une seule limite, dont l'Analyse mathématique s'est seulement intéressée jusqu'ici, ne constitue pas un résultat particulièrement nouveau, mais c'est un point de vue pouvant donner le ressort à un nouveau chapitre de l'Analyse ; les progrès plus ou moins récents du Calcul des Probabilités autorisent à croire que ces études peuvent être susceptibles de très intéressants développements.

LA CONSTRUCTION DU CHAMP DE PROBABILITÉ AVEC LA SOLUTION FONDAMENTALE DE L'ÉQUATION PARABOLIQUE NORMALE AUX CŒFFICIENTS HOLDERIENS

Halina MILICER GRUŻEWSKA

INTRODUCTION

On a démontré que la solution de l'équation parabolique normale ((1. 1) [1]), sans terme linéaire, dite l'équation de Kolmogoroff, qui n'est autre que la solution fondamentale normée, est la densité de la probabilité de passage (fonction de passage) de processus stochastique. Les hypothèses qu'on y fait se trouvent précisées dans l'article [1] comme Supposition I et II, et la solution en question c'est la fonction $U(X, t. Y, \tau)$, définie par la formule (1. 2) du même article. On a démontré aussi [2], qu'on peut définir dans chaque champ de probabilité la probabilité conditionnelle (ou de passage). Dans cet article au contraire on demande, si l'on peut construire le champ de probabilité, dès qu'on a la solution de la dite équation (1. 1), [1] : $U(X, t; Y, \tau)$. La réponse est positive.

1/ On déduit facilement des articles [1] et [2], que la fonction :

$$f(X, t, Y, \tau) = f(X, t)\, U(X, t; Y\, \tau) \qquad (1.\ 1)$$

- dite fonction de champ dans R_{2n} - sera la densité de la probabilité, définie pour $X \subset R_n$, $Y \subset R_n$, $O \leqslant t < \tau \leqslant T$ = const, si seulement la fonction $f(X, t)$ est la densité de la répartition définie pour $X \in R_n$, $O \leqslant t \leqslant T$ - dite répartition marginale horizontale - Posons :

$$\varphi(Y, t, \tau) = \int_{R_n} f(u, t; Y, \tau)\, du = \int_{R_n} f(u, t)\, U(u, t; Y, \tau)\, du > 0, \qquad (2.\ 1)$$

$$t < \tau, \quad Y \in R_n.$$

Il nous faut savoir que $\varphi(Y, t, \tau)$ est une fonction de champ dans R_n. Mais c'est une conséquence immédiate de la définition de la fonction $U(X, t; Y, \tau)$, [1] et de l'expression (2. 1). On sait aussi, que :

$$\lim_{t \to T} \varphi(Y, t, \tau) = f(Y, \tau), \qquad (3.\ 1)$$

dès que la fonction $f(X, \tau)$ est continue au point (Y, τ), [3], [1]. Elle définit aussi la répartition marginale verticale, [2], par la formule :

$$\int_{-\infty}^{y} \varphi(v, t, \tau)\, dv = F(Y, t, \tau), \quad Y \in R_n, \; 0 \leqslant t \leqslant \tau \leqslant T. \qquad (3'.1)$$

$\varphi(Y, t, \tau)$ est donc la densité de $F(Y, t, \tau)$. La densité de la composante horizontale sera :

$$U^*(Y, \tau; X, t) = [\, f(X, t)/ \varphi(Y, t, \tau)\,] \,.\, U(X, t; Y, \tau); \qquad (4.1)$$

les deux composantes du champ verticale et horizontale seront donc :

$$\int_{-\infty}^{y} U(X, t; v, \tau)\, dv \quad \text{et} \quad \int_{-\infty}^{x} U^*(Y, \tau; u, t)\, du \qquad (5.1)$$

Soit maintenant $\bar{\psi}(u) = 0$ l'équation conjuguée à l'équation $\hat{\psi}(u) = 0$, dite équation (1.1), [1]. Nous savons, que si les coefficients de l'équation $\bar{\psi}(u) = 0$ vérifissent les conditions d'Hölder, alors :

$$\bar{\psi}_{y,\tau}(\Gamma(X, t; Y, \tau)) = 0. \qquad (6.1)$$

Il résulte alors de la définition (2.1) et de l'égalité (3.1), par la même méthode que dans l'article [1], que :

$$\varphi(Y, t, \tau) = f(Y, \tau), \qquad (7.1)$$

sous la supposition que la fonction : $f(Y, \tau)/A(Y, \tau)$ est la solution de l'équation $\bar{\psi}(u)$) = 0. En effet :

$$\varphi(Y, t, \tau) = \lambda \int_{Rn} f(u, t)\, A(Y, \tau)\, \Gamma(u, t; Y, \tau)\, du, \quad \text{donc :}$$

$$\varphi(Y, t, \tau)/A(Y, \tau) = \lambda \int_{Rn} f(u, t)\, \Gamma(u, t; Y, \tau)\, du \to f(Y, \tau)/A(Y, \tau),$$

parce que les coefficients des parties principales des équations conjuguées sont égaux. Outre cela on sait que :

$$\bar{\psi}_{y,\tau}\Big(\int_{Rn} f(u, t)\, \Gamma(u, t; Y, \tau)\, du\Big) = 0$$

Alors, si on a :

$$\bar{\psi}_{y,\tau}(f(Y, \tau)/A(Y, \tau)) = 0, \quad \text{alors :} \qquad (6'.1)$$

$$\lambda \int_{Rn} f(u, t)\, \Gamma(u, t; Y, \tau)\, du = f(Y, \tau)/A(Y, \tau), \quad \text{car :}$$

$$f(Y, t)/A(Y, t) \to f(Y, \tau)/A(Y, \tau), \quad t \to \tau.$$

On a donc démontré que :

$$\lambda \int_{Rn} f(u, t)\; A(Y, \tau)\, \Gamma(u, t; Y, \tau)\; du \equiv f(Y, \tau),$$

ou que :

$$\varphi(Y, t, \tau) \equiv f(Y, \tau), \text{ c. q. f. d...}$$

On a donc alors :

$$f(Y, \tau) = \int_{Rn} f(u, t)\; U(u, t; Y, \tau)\; du.$$

D'ici résulte pour la fonction de champ dans R_{2n} une formule élégante.

$$f(X, t\,; Y, \tau) = f(X, t)\, U(X, t.\, Y, \tau) = f(Y, \tau)\, U^{*}(Y, \tau\,.\, X, t). \quad (8.1)$$

Je nommerai dès lors le champ des mêmes probabilités marginales par :

<u>champ simple.</u>

On voit que dans ce cas les probabilités marginales du champ construit avec la solution fondamentale de l'équation (1. 1), [1], ne sont pas arbitraires : elles sont toutes les deux une solution de l'équation de Fredholm de seconde espèce :

$$f(Y, \tau) = \int_{Rn} U(u, t.\, Y, \tau)\, f(u, t)\, du, \qquad (9.1)$$

dont le noyau $U(X, t; Y, \tau)$ dépend des coefficients de l'équation (1.1), [1], sous les hypothèses I et II de l'article [1], et les suppositions (6. 1) et (6.1) de cet article. Mais la construction du champ dans le cas général, qui est définie par les formules de (1. 1) à (5. 1) repose seulement sur les hypothèses I et II, [1]. Ce champ pourrait être aussi simple, dès que nous pourrions affirmer que l'équation (9. 1) possède des solutions ou une solution. Une telle solution existerait, on le sait, si le domaine d'intégration dans l'intégrale de la formule (9. 1) serait borné. En résumant nous pouvons écrire : "La solution fondamentale de l'équation 1er de Kolmogoroff, sous les suppositions I et II de l'article [1], étant normalisée, alors la fonction $U(X, t\,; Y, \tau)$, ((1. 2), [1]), est une fonction de passage d'un champ de probabilité dans R_{2n}, défini par les formules de (1. 1) à (5. 1). Si la densité $f(X, t)$ de la probabilité marginale horizontale est arbitraire, elle peut être différente de la densité $\varphi(Y, t, \tau)$ de la probabilité marginale verticale, qui peut dépendre aussi du temps passé t. Si au contraire cette densité $f(X, t)$ est une solution de l'équation (9. 1), alors la densité $\varphi(Y, t, \tau)$ lui est égale. "En particulier on a le Théorème suivant :

THEOREME (1. 1) -

Pour que le champ de probabilité, construit avec la solution fondamentale normalisée de l'équation 1er de Kolmogoroff, assujetti aux

conditions I et II, [1], soit simple il faut et il suffit, que l'équation intégrale (9. 1) possède la solution.

2/ Pour résoudre l'équation (2. 1) faisons intervenir la seconde intégrale généralisée de Poisson-Weierstrass [3], [4], à savoir :

$$I_1(Y, t, \tau) = \int_{R_n} \rho(u, t)\, \Gamma(u, t; Y, \tau)\, du\,;\ t < \tau \qquad (1.2)$$

où $\rho(u, t)$ est une fonction intégrable et bornée dans R_n. C'est évident, qu'au cas où la supposition (6. 1) est valable on a $\bar{\Psi}_{y,\tau}(I_1(Y, t, \tau) \equiv 0$. On y remplace la fonction $\Gamma(u, t; Y, \tau)$, pour la normer, par la fonction $U(u, t; Y, \tau)$ et l'intégrale $I_1\ (Y, t, \tau)$ par $\bar{I}(Y, t, \tau)$, en écrivant :

$$\bar{I}(Y, t, \tau) = \lambda I_1(Y, t, \tau)\, A(Y, \tau) = \int_{R_n} \rho\, (u, t)\, U(u, t; Y, \tau)\, du. \qquad (2.2)$$

On démontre comme dans [3], [4], que si la fonction $\rho(u, t)$ est continue au point (Y, t), alors :

$$\lim_{\tau \to t} \bar{I}(Y, t, \tau) = \rho(Y, t). \qquad (3.2)$$

Nommons par $f(u, O)$ une fonction continue du champ dans R_n, et la fonction $\bar{I}\ (Y, O, \tau)$, qui lui correspond, par $f(Y, \tau)$, de sorte que :

$$f(Y, \tau) = \int_{R_n} f(u, O)\ U(u, O; Y, \tau)\, du. \qquad (4.2)$$

Mais nous savons déjà que :

$$f(Y, \tau) \to f(Y, O),\ \ \tau \to O,\ \ f(Y, \tau) > 0,\ \ \int_{R_n} f(Y, \tau)\, dY = 1. \quad (5.2)$$

La fonction $f(Y, \tau)$ est donc une fonction de champ. On démontrera, que c'est la solution de l'équation intégrale (9. 1). En effet, on reçoit, en la substituant dans la partie droite de cette équation, après avoir changé la variable d'intégration u par v dans (4. 2) :

$$\int_{R_n} U(u, t; Y, \tau)\, f(u, t)\, du = \int_{R_n} U(u, t; Y, \tau) \int_{R_n} f(v, O).\ U(v, O, u, t)\, dv\, du;$$

D'où on reçoit par le changement de l'ordre d'intégration :

$$\int_{R_n} U(u, t; Y, \tau)\, f(u, t)\, du = \int_{R_n} f(v, O) \int_{R_n} U(v, O; u, t)\, U(u, t; Y, \tau)\, du\, dv$$

ce qui donne d'après les résultats de l'article [1] :

$$\int_{R_n} U(u, t; Y, \tau)\ f(u, t)\ du = \int_{R_n} f(v, O)\ U(v, O; Y, \tau)\, dv,$$

d'ici et de la définition (1. 2) résulte l'égalité (9. 1), c. q. f. d... On a donc démontré le :

THEOREME (1. 2) -

La solution non nulle de l'équation de Fredholm (9. 1) est la seconde intégrale généralisée de Poisson-Weierstrass , de la solution fondamentale normée du 1er équation de Kolmogoroff, pour t = 0, et d'une fonction de champ dans R_n, qui est continue. L'équation est assujetie aux suppositions I et II, [1]. D'ici résulte comme corollaire le :

THEOREME (2. 2) -

Admettons les hypothèses I et II de l'article [1]. La solution fondamentale, normée $U(X, t. Y, \tau)$ de la première équation de Kolmogoroff $\hat{\psi}(U) = O$, et la seconde intégrale généralisée de Poisson - Weierstrass, de cette solution pour t = 0, et d'une fonction de champ, continue dans R_n, construisent dans R_{2n} le champ simple. Ces deux composantes vérifient l'équation de Smoluchowski

En effet. Il résulte de l'article [1] et des résultats précédents (à savoir des formules (4.2) et (5. 2)), qu'il ne suffit de démontrer l'égalité de Smoluchowski que pour la densité $U^*(Y, \tau ; X, t), \tau > t$ de la composante horizontale. Elle est définie par les formules (4. 1), (2. 1) et (9. 1), d'où résulte que :

$$U^*(Y, \tau ; X, t) = [f(X, t) / f(Y, \tau)] \, . \, U(X, t ; Y, \tau).$$

où $f(Y, \tau)$ est définie par l'égalité (4. 2). On a donc :

$$\int_{R_n} U^*(Y, \tau ; Z, \theta) \, U^*(Z, \theta ; X, t) \, dZ =$$

$$\int_{R_n} [f(Z, \theta) / f(Y, \tau)] \, . \, U(Z, \theta ; Y, \tau) \, [f(X, t) / f(Z, \theta)] \, . \, U(X, t; Z, \theta) \, dZ =$$

$$[f(X, t)/f(Y, \tau)] \, U(X, t; Y, \tau) = U^*(Y, \tau ; X, t), \text{ c. q. f. d...}$$

Remarque.

L'égalité (6'. 1) résulte des égalités (4. 2) et (6'. 1)

REFERENCES BIBLIOGRAPHIQUES

1) - H. MILICER GRUŻEWSKA, Comptes rendus, 247, 1958 p. p. 1168-1171.

2) - " " " , Comptes rendus, 232, 1951, p. 1256.

3) - W. POGORZELSKI, Annales Polonici Mathematici, IV, 1, 1957, p. p. 61-92.

4) - W. POGORZELSKI, Annales Polonici Mathematici, IV, 3, 1958, p. p. 292-297.

SUR LA SOLUTION DE L'ÉQUATION PARABOLIQUE NORMALE AUX CŒFFICIENTS HÖLDERIENS QUI EST UNE PROBABILITÉ CONDITIONNELLE DE PROCESSUS STOCHASTIQUE

PARTIE I

Halina MILICER GRUŻEWSKA

INTRODUCTION

On sait, certaines conditions étant admises, que la probabilité de passage de processus stochastique est une solution de la première équation de Kolmogoroff [1], et que la solution de deux équations de Kolmogoroff est une fonction de passage. Les solutions des équations, dont il s'agit, se trouvent chez Feller [2] et Dressel [3]. Mais la supposition de la continuité des secondes dérivées des coefficients de la forme caractéristique de la première équation de Kolmogoroff est essentielle.

Récemment l'équation parabolique fût résolue par M. W. Pogorzelski (voir [4] et [7]) sous l'hypothèse que les coefficients qui figurent dans cette équation sont hölderiens donc sous une hypothèse beaucoup plus générale. Dans ces conditions la seconde équation de Kolmogoroff peut ne pas exister. On verra, que le théorème d'unicité de la solution de l'équation parabolique [5] suffit pour démontrer, que la solution de la première équation de Kolmogoroff est la densité de la probabilité de passage.

Les notations sont appropriées de l'article [4].

1/ Soit l'équation parabolique normale avec un signe changé (près de u'_t).

$$\hat{\varphi}(u) = \sum_{i,j=1}^{n} a_{ij}(X, t)\, u''_{x_i x_j} + \sum_{i=1}^{n} a_i(X, t)\, u'_{x_i} + c(X, t)\, u + u'_t = 0^{(1)}. \quad (1.1)$$

Nous admettons les hypothèses suivantes :

-I. Les coefficients $a_{ij}(X, t)$, $a_i(X, t)$ et $c(X, t)$ sont définis, bornés

---

(1) Pour $c(x, t) \equiv 0$ c'est l'équation de Kolmogoroff.

et continus dans la région ($X \in R_n$, $0 \leqslant t \leqslant T$), R_n désignant l'espace euclidien des points $X(x_1, x_2, \dots x_n)$.

-II. ces coefficients sont soumis aux conditions d'Hölder, a savoir :

$$|a_{ij}(X, t_1) - a_{ij}(Y, t_2)| \leqslant k(|X-Y|^h + |t_1 - t_2|^{h'}), \ 0 < h, h' \leqslant 1,$$
$$|a_i(X, t) - a_i(Y, t)| \leqslant k' \, |X - Y|^h, \ 0 < h \leqslant 1, \tag{2. 1}$$
$$|c(X, t) - c(Y, t)| \leqslant k'' \, |X - Y|^h, \quad '' \quad , \text{ où :}$$

h, h', k, k' et k'' désignent des constantes positives, $Y = Y(y_1, \dots y_n)$ et

$$|X - Y|^2 = \sum_{i=1}^{n} (x_i - y_i) = r_{XY}^2 \ ;$$

-III. la forme quadratique $\sum_{ij} a_{ij} \lambda_i \lambda_j$: est définie positive et, si $a^{ij}(Z, t)$, $Z = Z(z, \dots z_n) \in R_n$, $0 \leqslant t \leqslant T$, désignant les éléments de la matrice inverse à la matrice $||a_{ij}(Z, t)||$, et que :

$$\delta^{Z,t}(X, Y) = \sum_{i,j=1}^{n} a^{ij}(Z, t)\,(x_i - y_i)\,(x_j - y_j)\,, \tag{3. 1}$$

alors il existe des telles constantes positives g et G que :

$$g|X-Y|^2 \leqslant \delta^{Z,t}(X, Y) \leqslant G|X-Y|^2. \ X, Y, Z \in R_n, 0 < t < T. \tag{4. 1}$$

En posant comme toujours:

$$w^{Z,\tau}(X, t; Y, \tau) = (\tau - t)^{-\frac{n}{2}} \exp\left[-\frac{\delta^{Z,\tau}(X, Y)}{4(\tau - t)}\right], \ \tau > t, \tag{5. 1}$$

on a comme solution fondamentale de l'équation (1. 1), (v.[4]) la fonction suivante :

$$\Gamma(X; t, Y, \tau) = w^{Y,t}(X, t; Y, \tau) + \int_t^{\tau}\int_{R_n} w^{Z,\theta}(X, t; Z, \theta)\Phi(Z, \theta; Y, \tau)\,dZ\,d\theta, \tag{6.1}$$

où $\Phi(X, t; Y, \tau)$ est la solution de l'équation intégrale :

$$\hat{\Psi}_{X,t}[w^{Y,\tau}(X, t; Y, \tau)] - (2\sqrt{\pi})^n A^{-1}(X, t)\,\Phi(X, t; Y, \tau) + \tag{7. 1}$$
$$+ \int_t^{\tau}\int_{R_n} \hat{\Psi}_{X,t}[w^{Z,\theta}(X, t; Z, \theta)\,\Phi(Z, \theta; Y, \tau)\,dZ\,d\theta = 0,$$

$A(X, t) = \sqrt{\det |a^{ij}(X, t)|}$.

En posant encore :

$$\lambda=(2\sqrt{\pi})^{-n};\ N(X,t;Z,\theta)=\lambda^{-1}f(X,t;Z,\theta)=\hat{\varphi}_{X,t}[w^{Z,\theta}(X,t;Z,\ )]\,A(X,t),\quad (8.1)$$

on peut ecrire que :

$$\Phi(X,t;Y,\tau)=f(X,t;Y,\tau)+\lambda\int_t^T\int_{R_n}\mathfrak{N}(X,t;Z,\theta)\,f(Z,\theta;Y,\tau)\,dZ\,d\theta,\ t<\tau \quad (9.1)$$

$\mathfrak{N}(X,t;Z,\theta)$ c'est le noyau résolvant de l'équation (7.1); les noyaux itérés du noyau $N(X,t;Y,\theta)$ sont déterminés par la relation de récurrence :

$$N_{n+1}(X,t;Y,\theta)=\int_t^\theta\int_{R_n}N(X,\theta;Z,\theta')\,N_n(Z,\theta';Y,\theta)\,dZ\,d\theta',\ N_o=N \quad (10.1)$$

L'existence des intégrales des équations (10.1), (9.1) et (7.1), pour $t<\theta\leqslant T$, résulte des articles [4] et[7].

M. Pogorzelski fait intervenir dans l'article [6] l'intégrale que voici :

$$J(X,t,\tau)=\int_{R_n}\Gamma(X,t;Y,\tau)\,\rho(Y,\tau)\,dY,\quad \text{ici}\quad t<\tau. \quad (11.1)$$

$\rho(Y,\tau)$ est une fonction bornée dans l'espace $(R_n,\ 0<\tau\leqslant T)$ et intégrable dans tout domaine mesurable et borné de cet espace. L'inté-(11.1) existe et vérifie l'équation (1.1) pour $0\leqslant t<\tau\leqslant T$.

Il fût démontré dans le même article [6], que si la fonction $\rho(X,\tau)$ est continue au point (X,t), alors on a la propriété limite :

$$\lim_{\tau\to t}J(X,t,\tau)=[\lambda\,A(X,t)]^{-1}\cdot\rho(X,t). \quad (12.1)$$

Rappelons que la fonction $\Gamma(X,t;Y,\tau)$ dans l'expression (11.1) est définie par les formules (6.1), ... (10.1).

2/ On fait intervenir l'expression que voici :

$$U(X,t;Y,\tau)=\lambda A(Y,\tau)\,\Gamma(X,t;Y,\tau),\ t<\tau\leqslant T. \quad (1.2)$$

Cette fonction est une solution de l'équation (1.1), pour $t<\tau$, par rapport à X et t.

Lemme (1.2).

Si la fonction $f(X,t)$ est une solution de l'équation (1.1), $X\in R_n$, $0\leqslant t\leqslant T$, ou elle est continue et bornée (voir [5]), alors on a l'identité suivante :

$$f(X,t) = \int_{R_n} U(X,t;Y;\tau)\, f(Y,\tau)\, dY, \quad 0 \leqslant t < \tau \leqslant T, \tag{2.2}$$

où la fonction $U(X, t; Y, \tau)$ est définie par l'égalité (1. 2), si les coefficients de l'équation (1. 1) vérifient les hypothèses I, II et III du 1/.

Preuve du Lemme (1. 2).

Nous savons, à cause des égalités (12. 1) et (1. 2), que :

$$\lim_{t \to \tau - 0} \int_{R_n} U(X, t; Y, \tau)\, f(Y,\tau)\, dY = f(X,\tau). \tag{3.2}$$

Mais la fonction $f(X, t)$ est la solution de l'équation (1. 1) et :

$$\lim_{t \to \tau} f(X, t) = f(X, \tau). \tag{4.2}$$

les fonctions $f(X, t)$ et l'intégrale :

$$\int_{R_n} U(X, t; Y, \tau)\, f(Y, \tau)\, dY ,$$

sont les solutions de la même équation (1. 1) avec les mêmes conditions finales : (3.2) et (4.2) et elles sont bornées, donc de classe E_2 (v. 5)(1), elles sont donc identiques, c. q. f. d. d. . .

On déduit immédiatement les corollaires suivants :

Corollaire (1. 2).

Si les coefficients de l'équation (1. 1) vérifient les hypothèses I, II et III du 1/ et que $c(X, t) = 0$, alors la fonction $U(X, t; Y, \tau)$, définie par l'expression (1. 2), satisfait à l'identité suivante :

$$\int_{R_n} U(X, t; Y, \tau)\, dY \equiv 1, \; X, Y \in R_n , \; t \leqslant \tau .$$

Preuve du Corollaire (1. 2).

On y pose dans l'expression (2. 2) : $f(Y, \tau) \equiv 1$, ce qui est légitime, parce que une telle fonction vérifie l'équation (1. 1), où on a posé $c(X, t) \equiv 0$, et $\lim f(X, t) \equiv 1$, $t \to \tau$, $X \in R_n$.

Corollaire (2. 2).

La fonction (1. 2) vérifie l'équation de Smoluchowski, à savoir :

$$U(X, t; Y, \tau) = \int_{R_n} U(X, t; Z, \theta)\, U(Z,\theta; Y, \tau)\, dZ \text{ pour : } t < \theta < \tau ,$$

si les conditions I, II et III du 1/ sont accomplies.

----

(1) La condition finale revient à la condition initiale par le changement de variable par exemple : $\bar{t}$ en $t = \tau - t$.

Preuve du Corollaire (2. 2).

On transforme l'expression (2.2) en écrivant : $Y = Z$ et $\tau = \theta > t$, et on y pose $f(Z, \theta) = U(Z, \theta; Y, \tau)$. C'est l'égitime parce que la fonction $U(X, t; Y, \tau)$ est une solution continue de l'équation (1. 1), pour $t \leqslant \theta < \tau$, et $\lim_{t \to \theta} U(X, t; Y, \tau) = U(X, \theta; Y, \tau)$. Outre celà la solution $U(X, t; Y, \tau)$ est bornée pour $0 \leqslant t \leqslant \theta < \tau$, θ étant fixé, d'après l'article [7], c. q. f. d. d.

On peut démontrer aussi (1), pareillement que l'a fait Feller dans l'article [2], que l'expression (1. 2) est non négative, sous les suppositions du Corollaire (1. 2). On a donc le Théorème suivant :

Théorème (1. 2).

La fonction (1. 2) est la densité de la probabilité de passage de l'état X, au moment t, à l'état Y, au moment $\tau > t$, $X, Y \in R_n$, si les coefficients de l'équation (1. 1) satisfont aux hypothèses du Corollaire (1. 2).

Les conditions de continuité de la fonction (1.2), à savoir l'existence des moments moyens (voir [2]), sera démontrée dans la seconde partie de cet article.

TRAVAUX CITES

(1) - A. KOLMOGOROFF, Mathematische Annalen, 194, 1931, p.p.

(2) - W. FELLER, Mathematische Annalen, 113, 1936, p.p. 113-160.

(3) - F.G. DRESSEL, DUKE, Mathematical Journal, 13, 1946, p.p. 61-70.

(4) - W. POGORZELSKI, Ricerche di Matematica, V, 1956, p.p. 25-57.

(5) - M. KRZYZANSKI, Annales Polonici Mathematici XVIII, 1945, p.p. 145-156.

(6) - W. POGORZELSKI, " " " IV, I, 1957, p.p. 61-92.

(7) - W. POGORZELSKI, " " " IV, 3, 1958, p.p. 288-307.

(8) - M. PICONE, Mathematische Annalen, 101, 1929, p.p. 701-712.

----

(1) Soit : $U(X', t'; Y', \tau') < 0$ et $c(X, t) \equiv 0$. Posons : $g(X, t) = U(X, t; Y', \tau')$ et $z(X, t) = g(X, t) \exp(-t+t')$. On a : $\hat{\Psi}_{x,t}(g(X, t)) \equiv 0$, $g(X', t') < 0$, alors $z(X, t)$ est une solution d'une équation parabolique normale avec le coefficient linéaire $c_1 = -1$; de plus $z(X', t') < 0$.
Soit : $t' < t_1 < \tau'$ et K un cercle de centre (X', t') et de rayon $t_1 - t'$. Sur la partie K_1 de ce cercle ou $t \geqslant t'$ la fonction $z(X, t)$ aura un minimum par exemple au point (X'', t'') et nous pouvons écrire, à cause du théorème de Gévrey, que $z(X'', t'') \leqslant z(X', t') < 0$. Mais c'est une contradiction avec le fait que $z(X'', t'')$ est la solution de la dite équation parabolique normale, où $c_1 = -1$, c. q. f. d. d.

SUR LA SOLUTION DE L'ÉQUATION PARABOLIQUE NORMALE AUX CŒFFICIENTS HOLDERIENS QUI EST UNE PROBABILITÉ CONDITIONNELLE DE PROCESSUS STOCHASTIQUE

PARTIE II

Halina MILICER GRUŻEWSKA

INTRODUCTION

On a démontré dans la première partie de cet article que la solution de la première équation de Kolmogoroff est une probabilité de passage, si les coefficients de cette équation sont soumis aux conditions I, II, III. Dans cette partie on prouvera, que cette solution accomplie les bien connues conditions de continuité de la probabilité de passage. Les notations de la partie première seront maintenues ici. Les dites conditions de continuité pourront donc être écrites comme il suit :

$$\tau - t = \Delta t > 0, \quad \mu_i(X, t, \tau) = \int_{R_n} (x_i - y_i)\, U(X, t; Y, \tau)\, dY = o(\sqrt{\Delta t}), \quad i = 1 \ldots n.$$

$$\lim_{\Delta t \to 0} (\Delta t)^{-1} \int_{R_n} (x_i - y_i)(x_i - y_i)\, U(X, t; Y, \tau)\, dY = 2\, a_{ij}(X, t), \quad i, j = 1 \ldots n.$$

$$\lim_{\Delta t \to 0} (\Delta t)^{-1} \int_{R_n} (y_i - x_i)\, U(X, t - \Delta t; Y, t)\, dY = b_i, \quad i = 1, \ldots, n.$$

Ces conditions de continuités, étant démontrées, permettront d'affirmer que la solution de la première équation de Kolmogoroff est non seulement une probabilité de passage, mais encore que ses deux premiers moments, centralisés vers l'état initial X, et divisés par l'accroissement du double de temps Δt, ou pour cet accroissement même, convergent vers les coefficients de la dite équation, quand $\Delta t \to 0$.

1/ - Lemme (1. 1).

Les premiers moments de la fonction de passage $U(X, t, Y, \tau)$, centralisés vers l'état initial X, sont $o(\sqrt{\tau - t})$, si les coefficients de la première équation de Kolmogoroff (équation (1. 1) dans la partie I de cet article) accomplissent les suppositions du Théorème (1. 2) de cette partie.

Nous pouvons écrire, avec les définitions (1. 2) p. I et de l'article (4) (p. I) que :

$$\mu_i(X, t, \tau) = \int_{R_n} (x_i - y_i)\, \rho(Y, \tau)\, [\, w^{Y,\tau}(X, t; Y, \tau) + \bar{w}(X, t; Y, \tau)]\, dY, \quad (1. 1)$$

où :

$\rho(Y,\tau) = \lambda A(Y,\tau)$

$w^{Y,\tau}(X, t; Y, \tau)$ est défini par l'équation (5. 1) p. I et

$\bar{w}(X, t; Y, \tau) = \int_t^{\tau} \int_{R_n} w^{Z,\theta}(X, t; Z, \theta)\, \Phi(Z, \theta, Y, \tau)\, dY\, d\theta$

v. (4) p. I, formules (98) et (99).

Nommons les intégrales correspondantes aux deux composants de l'expression (1. 1) par I_1 et I_2 respectivement. Nous avons donc :

$$\mu_i(X, t, \tau) = I_1 + I_2.$$

On transformera le premier composant comme il suit :

$$I_1 = \rho(X,\tau) \int_{R_n} (x_i - y_i)\, w^{X,t}(X, t; Y,\tau)\, dY + \qquad (2.\ 1)$$

$$\rho(X,\tau) \int_{R_n} (x_i - y_i)\, [w^{Y,\tau}(X, t; Y,\tau) - w^{X,t}(X, t, Y, \tau)]\, dY +$$

$$\int_{R_n} [\rho(Y,\tau) - \rho(X,\tau)]\, (x_i - y_i)\, w^{Y,\tau}(X, t; Y, \tau)\, dY.$$

Mais il résulte de la définition de la fonction $\rho(Y,\tau)$ et des suppositions I, II, III p. I que :

$$\rho(Y,\tau) - \rho(X,\tau) = 0\,(r_{XY}^{h/2}), \qquad (3.\ 1)$$

et que :

$$w^{Y,\tau}(X, t; Y, \tau) - w^{X,t}(X, t; Y, \tau) = 0 \left\{ n^2 r_{XY}^2 (\tau - t)^{-\frac{n}{2}-1} \sup |a^{ij}(Y,\tau) - a^{ij}(X, t)| \right.$$

$$\left. \exp\left[- \frac{g\, r_{xy}^2}{\tau - t} \right] \right\}.$$

v. p. ex. (4) p. I formule (16).

Outre celà on reçoit par une transformation linéaire convenable de la variable Y l'égalité suivante :

$$\int_{R_n} (y_i - x_i)\, w^{X,t}(X, t; Y,\tau)\, dY =$$

$$(\tau - t)^{-\frac{n}{2}} \int_{R_n} \sum_{\nu=1}^{n} c_\nu^i z_\nu \exp\left[-\sum_{\nu=1}^{n} z_\nu^2 / 4(\tau - t) \right] dz_1, \ldots, dz_n \,.\, A^{-1}(X, t) \equiv 0.$$

où c_ν^i sont des constantes.

Ces égalités permettent d'écrire que :

$$I_1 / \sqrt{\tau - t} = 0\left\{ \int_{R_n} (x_i - y_i)(\tau - t)^{-\frac{n}{2}-1\frac{1}{2}} . r_{XY}^2 \sup | a^{ij}(Y,\tau) - \right.$$

$$\left. - a^{ij}(X,t) | \exp[-g r_{XY}^2 / (\tau - t)] \, dY \right\} +$$

$$+ 0\left\{ \int_{R_n} (x_i - y_i)(\tau - t)^{-\frac{n+1}{2}} r_{XY}^{h/2} \exp[-g r_{XY}^2 / (\tau - t)] \, dY \right\}.$$

Chacune de ces intégrales est représentée comme une somme des deux intégrales répendues respectivement sur les domaines : $r_{XY} \leqslant 1$ et $r_{XY} > 1$. Cette dernière converge vers zéro avec $(\tau - t)$, comme il est facile de voir.

Quant à la première on y pose :

$$r_{XY} / \sqrt{\tau - t} = u,$$

d'où on reçoit à cause de la supposition II. p. I, l'inégalité suivante :

$$\sup | a^{ij}(Y,\tau) - a^{ij}(X,t) | \leqslant k[(u\sqrt{\tau - t})^h + (\tau - t)^{h'}] ,$$

et on peut écrire :

$$I_1 = 0(1) + 0\{ \int_{|u| \leqslant 1/\sqrt{\tau - t}} u^{n+2-h} [u(\tau - t)^{h/2} + (\tau - t)^{h'}] \exp(-gu^2) \, du \} +$$

$$0\left\{ \int_{|u| \leqslant 1/\sqrt{\tau - t}} u^{n+h/2} (\tau - t)^{h/4} \exp(-gu^2) \, du \right\} \to 0, \ (\tau - t) \to 0 , \quad \text{c. q. f. d. d .}$$

L'intégrale I_2 sera traitée comme dans l'article (4) p. I. à savoir :

$$I_2 = \int_{R_n} \rho(Y,\tau)(x_i - y_i) \bar{w}(X,t;Y,\tau) \, dY = \int_t^\tau \int_{R_n} w^{Z,\theta}(X,t;Z,\theta) \bar{\rho}_i(Z,\theta,\tau) \, dZ \, d\theta,$$

où :

$$\bar{\rho}_i(Z,\theta,\tau) = \int_{R_n} \Phi(Z,\theta,Y,\tau) \rho(Y,\tau)(x_i - y_i) \, dY.$$

Posons maintenant :

$$x_i - y_i = x_i - z_i + z_i - y_i ,$$

et écrivons comme dans les articles (4) et (7) p. I que :

$$(x_i - y_i) \Phi(Z,\theta,Y,\tau) = 0\{ r_{ZY}^{-n_1} . (\tau - \theta)^{-\mu_1} \} + 0\{ r_{ZY}^{-n_2} . (x_i - z_i) \} ,$$

où :

$$n_1 = n + 1 - 2\mu_1 - h_1 , \quad n_2 = n + 2 - 2\mu_1' - h_1$$

et :

$$1/2 - h_1/2 < \mu_1 < 1/2, \text{ pour } r_{ZY} \leqslant 1 \quad \text{et } \mu_1' > 1 - \frac{h_1}{2}.$$

$$\mu_1 < 1/2 - h_1, 2, \text{ pour } r_{ZY} > 1 , \quad \text{et} \quad \mu_1' < 1 - \frac{h_1}{2},$$

c. à. d.

$$\mu_1 = 1/2 - h_1 \ 2 \ \pm \varepsilon_1 , \ \mu_1' = 1 - h_1 2 \pm \varepsilon_1' ,$$

où ε_1, $\varepsilon_1' > 0$ seront petits.

Il résulte d'ici, que :

$$\bar{\rho}_i (Z, \theta, \tau) = 0 \{ (\tau - \theta)^{-\mu_1} + (x_i - z_i)(\tau - \theta)^{-\mu_1'} \} .$$

On aura alors, pour $O < \mu$, mais $\mu + \mu_1 < 1/2$ et $\mu' = -1/2 \pm \varepsilon'$, $\varepsilon' > O$, mais $\varepsilon' + \varepsilon_1' < h_1/2$, l'égalité suivante :

$$I_2 = 0 \{ \int_t^{\tau} [(\theta - t)^{-\mu} . (\tau - \theta)^{-\mu_1} + (\theta - t)^{-\mu'} (\varepsilon - \theta)^{-\mu_1'}] d\theta \} =$$

$$0 [(\tau - t)^{1-\mu-\mu_1} + (\tau - t)^{1-\mu'-\mu_1'}] = 0 [(\tau - t)^{1/2}] ,$$

comme $\mu' + \mu_1' < 1/2$, c. q. f. d. d...

2/ - Lemme (1. 2).

Les valeurs moyennes des secondes, moments centralisés vers l'état initial X, de la probabilité de passage U (X, t, Y, τ), qui est la solution de l'équation (1. 1) p. I, tendent vers les doubles des coefficients respectifs de la forme quadratique de cette équation avec $(\tau - t) \to 0$, si les suppositions du Théorème (1. 2) p. I. sont vérifiées. A savoir :

$$\lim_{\tau - t \to 0} (\tau - t)^{-1} \int_{R_n} (x_i - y_i)(x_j - y_j) U(X, t; Y, \tau) dY = 2 a_{ij}(X, t), \quad i, j = 1, 2, .. n. \quad (1.2)$$

Preuve du Lemme (1. 2).

On écrit, conformément aux notations qui précèdent :

$$(\tau - t)^{-1} \int_{R_n} (x_i - y_i)(x_j - y_j) U(X, t; Y, \tau) dY =$$

$$(\tau - t)^{-1} \int_{R_n} (x_i - y_i)(x_j - y_j) [w^{Y,\tau}(X, t; Y, \tau) + \bar{w}(X, t; Y, \tau)] \rho(Y, \tau) dY = \quad (2.2)$$

$$(\tau - t)^{-1} \rho(X, t) \int_{R_n} (x_i - y_i)(x_j - y_j) w^{X,t}(X, t; Y, \tau) dY + 0(1) .$$

La dernière égalité c'est une conséquence des formules (3. 1) et de l'analogue à (2. 1) [1].

Désignons maintenant la dernière intégrale de l'égalité (2. 2) par B_{ij} et formons la somme qui suit :

$$\sum_{i,j} a^{ij}(X,t)\, B_{ij} = \beta_n \int_{R_n} \sum_{i,j} a^{ij}(X,t)\, u_i u_j \exp[-\delta^{X,t}(U,O)]dU. \quad (3.2)$$

où on a posé :

$$-x_i + y_i = 2u\sqrt{\tau - t}, \quad U = (u_1, \dots u_n) = \begin{Bmatrix} u_1 \\ u_2 \\ u_n \end{Bmatrix} \text{ la matrice d'une colonne},$$

$$\delta^{X,t}(U,O) = \sum_{i,j} a^{ij}(X,t)\, u_i u_j ,$$

$$\beta_n = (\tau - t)^{-\frac{n}{2}-1}\, 2^{n+2}\, \rho(X,t)\, (\tau - t)^{\frac{n}{2}+1} = 2^{n+2}\, \rho(X,t) = 2^2 \pi^{-\frac{n}{2}}. A(X,t)$$

Mais nous pouvons écrire :

$$\| a^{ij}(X,t) \| = \| a_{ij}(X,t) \|^{-1} = A^{-1} = A_1 \quad \text{et} \quad \delta^{X,t}(U,O) = U' A_1 U ,$$

en usant la notiation U' pour la matrice oposée à U.et pour la matrice inverse à la matrice A la notiation A^{-} . Soit maintenant la matrice orthogonale et normée C, et soit :

$$U = CV, \quad \text{donc } U' = V' C'$$

$$U' A^{-1} U = V' C' A^{-1} CV = V' K V , \quad \text{avec } K = C' A_1 C = \begin{Bmatrix} k_1 & \dots & 0 \\ 0 & k_2 \dots & 0 \\ 0 & 0 \dots & k_n \end{Bmatrix}$$

On peut donc écrire au lieu de la somme (3.2) l'égalité :

$$\sum_{i,j} a^{ij}(X,t)\, B_{ij} = \beta_n \int_{R_n} V' K V . \exp(-V'KV)\, dV \quad (4.2)$$

parce que dU = dV.

Mais on sait que :

$$V' K V = \sum_{i=1}^{n} k_i v_i^2 , \quad \text{si } V = \{ v_i , v_2 , \dots v_n \}$$

et alors on a : au lieu de l'égalité (4.2) l'égalité suivante :

$$\sum_{i,j} a^{ij}(X,t)\, B_{ij} = \beta_n \sum_{i=1}^{n} \int_{R_n} k_i v_i^2 \exp.(-\sum k_i v_i^2)\, dV. \quad (5.2)$$

Changeons les variables d'intégration en posant :

$$\sqrt{k_i}\, v_i = z_i , \quad \text{alors } dV = dZ/\sqrt{k_1 k_2 \dots k_n} = dZ/\sqrt{|A_1|}.$$

On reçoit donc :

$$\sum_{i,j} a^{ij}(X, t)\ B_{ij} = 4\pi^{-n/2} \sum_{i} \int_{R_n} z_i^2 \exp\left(-\sum_{i=1}^{n} z_i^2\right) dZ = 4\pi^{-\frac{n}{2}}\ 2^{-1}\ \pi^{\frac{n}{2}} n = 2n.$$

Mais on peut écrire que :

$$\sum_{i,j} a^{ij}(X,t)\ B_{ij} = \sum_{j=1}^{n} (a^{1j}(X, t)\ B_{1j} + \ldots\ a^{nj}(X, t)\ B_{nj}) = 2\sum_{j=1}^{n} 1$$

Les sommes des produits :

$$\alpha_{jj} = a^{1j}(X, t)\ B_{1j} + \ldots\ a^{nj}(X, t)\ B_{nj}$$

ce sont ces éléments du produit des déterminents A_1 et B qui se trouvent sur la diagonale principale. Soit le déterminent d'ordre n de la forme :

$$K_1 = \begin{vmatrix} 2, & 0, & \ldots & 0 \\ 0, & 2, & \ldots & 0 \\ 0, & 0, & \ldots & 2 \end{vmatrix} = 2^n$$

et soit C_1 un déterminent qui donne l'égalité qui suit :

$$C_1'(A_1 B)\ C_1 = \begin{vmatrix} \alpha_{11} & 0 & \ldots & 0 \\ 0 & \alpha_{22}, & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0, & 0, & \ldots & \alpha_{nn} \end{vmatrix}$$

alors on aura :

$$\{1\}C_1'(A_1 B)\ C_1(1) = \sum_{j=1}^{n} \alpha_{jj}\ ;\ \{1\}\ K(1) = \sum_{j=1}^{n} 2,$$

d'où :

$\{1\}C_1'(A_1 B)C_1(1) = \{1\}K_1(1)$, donc $C_1' A_1 B C_1 = K_1$, et $A_1 B = C_1 K_1 C_1' = 2^n$

parce que : $C_1 K_1 C_1' = 2^n C_1 C_1' = 2^n$.

Mais on a :

$$A_1^{-1}\ A_1 B = A_1^{-1}\ .\ 2^n \text{ et } B = A_1^{-1}\ .\ 2^n = 2^n\ .\ A, \text{ c. à d.}$$

on a :

$$B_{ij} = 2\ a_{ij}(X, t),$$

ce qui prouve, avec l'égalité (2. 2), le Lemme (1. 2).

3/ - Lemme (1. 3).

Les valeurs moyennes des premiers moments, centralisés vers l'état initial X, de la probabilité de passage $U(X, t, \tau)$, solution de l'é-

quation (1.1) p. I, tendent vers les coefficients des dérivées premières de cette équation, si les suppositions du Théorème (1.2) p. I sont vérifiées, avec $(\tau - t) \to 0$. A savoir :

$$\lim_{\tau - t \to 0} (\tau - t)^{-1} \int_{R_n} (y_i - x_i)\, U(X, t ; Y, \tau)\, dY = b_i(X, t), \quad i = 1, .. n, \quad (1.3)$$

Preuve du Lemme (1. 3).

La démonstration est ici classique (v. p. ex. (2) p. I). On part de l'équation de Smoluchowski pour la fonction de passage $U(X, t-\Delta t; Z, t)$ et $t - \Delta t < t < \tau$:

$$U(X, t-\Delta t, Y, \tau) = \int_{R_n} U(X, t-\Delta t ; Z, t)\, U(Z, t, Y, \tau)\, dZ,$$

mais d'ici :

$$U(X, t-\Delta t, Y, \tau) = U(X, t, Y, \tau) + \sum_{i=1}^{n} U'_{x_i}(X, t; Y, \tau) \int_{R_n} (z_i - x_i)\, U(X, t-\Delta t, Z, t)\, dZ +$$

$$\frac{1}{2} \sum_{i, j=1}^{n} U''_{x_i x_j}(X, t, Y, \tau) \int_{R_n} (x_i - z_i)(x_j - z_j)\, U(X, t-\Delta t ; Z, t)\, dZ + 0(1), \quad X \neq Y$$

donc :

$$0 = U'_t + \sum_{t=i}^{n} U'_{x_i} \lim_{\Delta t \to 0} 1/\Delta t \int_{R_n} (z_i - x_i)\, U(X, t-\Delta t; Z, t)\, dZ + X \neq Y + \sum_{i=j}^{n} U''_{x_i x_j} a_{ij}(X, t).$$

Mais la fonction U est la solution de l'équation (1.1) p. I, avec le coefficient $c(X, t) \equiv 0$. D'ici et de l'égalité précédente résulte que :

$$\lim_{\Delta t \to 0} 1/\Delta t \int_{R_n} (z_i - x_i)\, U(X, t - \Delta t; Z, t)\, dZ = b_i, \quad i = 1, \ldots n, \text{ c.q.f.d. d.}$$

Le Théorème (1. 2) p. I, les Lemmes (1. 1), (1. 2) et (1. 3) p. II permettent d'énoncer le Théorème (1. 3) :

THEOREME (1. 3) -

Les suppositions du Théorème (1. 2), étant admises, la solution de l'équation (1.1) p. I est la fonction de passage de l'état X au moment t à l'état Y au moment τ, et les coefficients de cette équation sont respectivement les valeurs limites moyennes de $\frac{1}{2}$ des moments seconds et des premiers, pour $(\tau - t) \to 0$. On peut traiter le coefficient du terme linéaire, qui est ici égal à zéro, comme limite moyenne du carré du premier moment, pour $(\tau - t) \to 0$. Tous les moments sont centralisés vers l'état initial X.

note de la page 9 -

(1) $(\tau - t)^{-1} \int_{R_n} \rho_1(Y,\tau)\, \bar{w}(X, t; Y, \tau)\, dY = d, \quad \rho_1(Y,\tau) = (x_i - y_i)(x_i - y_i)\, \rho(Y,\tau)$

donc :

$$d = (\tau - t)^{-1} \int_t^\tau \int_{R_n} w^{z,\theta}(X, t; Z, \theta)\, \bar{\rho}_1(Z, \theta, \tau)\, dZ\, d\theta$$, où : si $|X - Y| = r_{xy}$,

alors :

$$\bar{\rho}_1(Z, \theta, \tau) = \int_{R_n} \Phi(Z, \theta; Y, \tau)\, \rho_1(Y,\tau)\, dY = 0 \left\{ \int_{R_n} r_{ZX}^{-n_1} (\tau - \theta)^{-\mu_1} (r_{ZY} + r_{ZX})^2\, dY \right\},$$

où : $n_1 = n + 2 - 2\mu_1 - h_1$.

On a donc :

$$d = 0 \left\{ \int_{R_n} r_{ZY}^{-n_1+2} (\tau - \theta)^{-\mu_1}\, dY + \int_{R_n} r_{ZX}^{-n_1+1} (\tau - \theta)^{-\mu_1'} r_{ZX}\, dY + \int_{R_n} r_{ZY}^{-n_1} r_{ZX}^2 (\tau - \theta)^{-\mu_1''} dY \right\}.$$

Il faudra poser $\mu_1'' = 1 - h_1/2 \pm \varepsilon_1''$ selon le cas ou $r_{ZY} \leqslant 1$, ou $r_{ZY} > 1$,
$\mu_1' = 1/2 - h_1/2 \pm \varepsilon_1'$ " " " " " , " " ,
$\mu_1 = - h_1/2 \pm \varepsilon_1$ " " " " " , " " .

En intégrant l'expression d il faudra accepter pour l'exposant de $(\theta - t)$ de la fonction dominante du quasipotentiel : $\mu'' = -1 \pm \varepsilon''$, $\mu' = -1/2 \pm \varepsilon'$, $\mu = \pm \varepsilon$, le signe + ou - étant assujeti au cas de $r_{ZX} \leqslant 1$ ou $r_{ZX} > 1$, on recevra ainsi comme l'exposant de $(\tau - t)$ dans l'expression d : $h_1/2 \pm \varepsilon_1 \pm \varepsilon$, ou $h_1/2 \pm \varepsilon_1' \pm \varepsilon'$, ou $h_1/2 \pm \varepsilon_1 \pm \varepsilon''$. Ils seront tous positifs dès que $\varepsilon + \varepsilon_1 < h_1/2$, $\varepsilon'' + \varepsilon_1' < h_1/2$ et $\varepsilon'' + \varepsilon'' < h_1/2$, ce qui est toujours possible. L'expression d sera donc 0 (1), c. q. f. d. d.

SUR LE DOMAINE D'ATTRACTION DE LA LOI DE POISSON

A. FUCHS ET N. ROBY

I - POSITION DU PROBLEME -

Pour tout entier $n > 0$ on considère n variables aléatoires (v.a) indicatrices indépendantes X_{kn} $(k = 1, .., n)$ où X_{kn} peut prendre les valeurs $\begin{cases} 1 & \text{avec probabilité } p_{kn} \\ 0 & \text{" \quad " \quad } 1-p_{kn} \end{cases}$ où $0 < p_{kn} < 1$. On posera

$$S_n = \sum_{k=1}^{n} X_{kn}, \qquad a_n = \sum_{k=1}^{n} p_{kn} = E\{S_n\}, \qquad \alpha_n = \operatorname{Max}_{k=1}^{n} p_{kn} \quad (0<\alpha_n<1)$$

et l'on étudiera le rôle des deux conditions suivantes :

$$\lim_{n \to \infty} a_n = a, \qquad 0 < a < \infty \tag{1}$$

$$\lim_{n \to \infty} \alpha_n = 0 \tag{2}$$

Pour tout nombre $a > 0$ désignons par Y_a la v.a. de Poisson du paramètre a, définie par $P\{Y_a = k\} = e^{-a} \frac{a^k}{k!}$ (k : entier ≥ 0). On sait que si (1) et (2) sont réalisées la suite de v.a. S_n tend en loi, lorsque $n \to \infty$, vers la v.a. Y_a.

Nous nous proposons, au § 2, de donner une démonstration simple de cette proposition. Au § 3 nous montrerons que, réciproquement, les conditions (1) et (2) sont nécessaires pour que la suite de v.a. S_n tende en loi, lorsque $n \to \infty$, vers la v.a. Y_a.

II - THEOREME DIRECT -

THEOREME I -

Si les conditions (1) et (2) sont réalisées, $S_n \longrightarrow Y_a$ en loi. La fonction caractéristique (f.c.) du S_n est :

$$\varphi_{s_n}(t) = \prod_{k=1}^{n} [1 + p_{kn}(e^{it} - 1)] , \qquad t \in R$$

$$\Longrightarrow \quad \Psi_{s_n}(t) = \text{Log } \varphi_{s_n}(t) = \sum_{k=1}^{n} \text{Log } [1 + p_{kn}(e^{it} - 1)]$$

d'où, en posant $e^{it} - 1 = z$:

$$\Psi_{s_n}(t) = \sum_{h=1}^{n} \text{Log } (1 + p_{kn} z)$$

Soit $\mathcal{V}$ un voisinage réel de $t = 0$ tel que $t \in \mathcal{V} \Longrightarrow |e^{it} - 1| = |z| < 1$. Alors $|p_{kn} z| < p_{kn} < 1$ et l'on peut développer $\text{Log } (1 + p_{kn} z)$ en série entière absolument et uniformément convergente dans $|z| < 1$.

$$\text{Log } (1 + p_{kn} z) = p_{kn} z + R_{kn}(z)$$

où $\displaystyle R_{kn}(z) = -\frac{(p_{kn} z)^2}{2} + \frac{(p_{kn} z)^3}{3} - \dots$

$$\Rightarrow \ |R_{kn}(z)| \leqslant \frac{p_{kn}^2 |z|^2}{2} + \frac{p_{kn}^3 |z|^3}{3} + \dots < \frac{p_{kn}^2}{2} + \frac{p_{kn}^3}{3} + \dots$$

$$\leqslant p_{kn}\left(\frac{\alpha_n}{2} + \frac{\alpha_n^2}{3} + \dots\right) = p_{kn} \frac{-\text{Log } (1 - \alpha_n) - \alpha_n}{\alpha_n} \text{ (car } 0 < \alpha_n < 1)$$

$$\Rightarrow \ \Psi_{s_n}(t) = a_n z + \sum_{k=1}^{n} R_{kn}(z)$$

$$\Rightarrow \ |\Psi_{s_n}(t) - a z| \leqslant |a_n z - a z| + \sum_{k=1}^{n} |R_{kn}(z)|$$

$$\text{Or} \begin{cases} |a_n z - a z| \leqslant |a_n - a| \, |z| < |a_n - a| & \forall \ z : |z| < 1 \\ \displaystyle\sum_{k=1}^{n} |R_{kn}(z)| \leqslant a_n \frac{-\text{Log } (1 - \alpha_n) - \alpha_n}{\alpha_n} & \forall \ z : |z| < 1 \end{cases}$$

On a donc, $\forall \ t \in \mathcal{V}$:

$$|\Psi_{s_n}(t) - a\,(e^{it} - 1)| \leqslant |a_n - a| + a_n \frac{-\operatorname{Log}(1-\alpha_n) - \alpha_n}{\alpha_n}$$

Or, en vertu de (1) et (2) le second membre tend vers 0 lorsque $n \longrightarrow \infty$. Mais puisque ce second membre majore le premier uniformément p.r. à $t \in \mathcal{V}$, il en résulte que le premier membre tend vers 0 lorsque $n \longrightarrow \infty$ et ceci uniformément p.r. à $t \in \mathcal{V}$. $\Longrightarrow$ $\Psi_{s_n}(t)$ tend vers $a\,(e^{it} - 1)$ lorsque $n \longrightarrow \infty$, cette convergence étant uniforme dans un voisinage réel $\mathcal{V}$ de $t = 0$. La conclusion en découle d'après un théorème classique de P. Lévy en remarquant que $a\,(e^{it} - 1)$ est la deuxième f.c. de la v.a. de Poisson Y_a.

III - THEOREME RECIPROQUE -

THEOREME III -

Si $S_n \longrightarrow Y_a$ en loi, alors les conditions (1) et (2) sont réalisées.

Supposons à présent que $S_n \longrightarrow Y_a$ en loi ; il en résulte d'après un théorème classique de P. Lévy, que sur tout intervalle réel borné de valeurs de t, en particulier pour $-\pi \leqslant t \leqslant +\pi$, la f.c. $\varphi_{s_n}(t) = \prod_{k=1}^{n} [1 + p_{kn}(e^{it} - 1)]$ de S_n tend uniformément vers la f.c. $e^{a(e^{it}-1)}$ de Y_a. C'est dire que le polynôme de la variable complexe u, $P_n(u) = \prod_{k=1}^{n} [1 - p_{kn} + p_{kn}\,u]$ tend uniformément vers la fonction holomorphe $e^{a(u-1)}$ sur le cercle $|u| = 1$, donc aussi dans le disque $|u| < 1$. En particulier, pour $u = 0$, $\prod_{k=1}^{n} (1 - p_{kn}) \longrightarrow e^{-a}$; pour n suffisamment grand on aura donc $e^{-2a} \leqslant \prod_{k=1}^{n} (1 - p_{kn})$; mais d'autre part $\prod_{k=1}^{n} (1 - p_{kn}) \leqslant e^{-a_n}$. Donc pour n suffisamment grand $e^{-2a} \leqslant e^{-a_n} \Longrightarrow a_n \leqslant 2a$. Il en résulte que a_n est borné quel que soit n :

$$a_n \leqslant A < \infty \qquad \forall n \text{ entier} > 0$$

Soit alors $\mathcal{C}(O,R)$ le disque du centre O et de rayon $R > 1$. Pour $u \in \mathcal{C}(O,R)$ on a $|P_n(u)| \leqslant \prod_{k=1}^{n} (1 - p_{kn} + p_{kn}\,R) \leqslant e^{a_n(R-1)} \leqslant e^{A(R-1)}$.

Les polynômes $P_n(u)$ sont donc uniformément bornés dans $\mathcal{C}(O,R)$; comme ils convergent uniformément vers $e^{a(u-1)}$ dans le disque $|u| < 1$ intérieur à $\mathcal{C}(O,R)$, il résulte du théorème de Vitali qu'ils convergent aussi uniformément vers la même fonction dans tout le disque $\mathcal{C}(O,R)$, en particulier dans tout un voisinage de

$u = 1$. Si l'on pose alors $u - 1 = z$, le polynôme $F_n(z) = \prod_{k=1}^{n} (1 + p_{kn} z)$ tend donc uniformément vers la fonction e^{az} dans tout un voisinage de $z = 0$: d'où convergence l'un vers l'autre des coefficients de Taylor de même rang.

En considérant les coefficients de z il vient ainsi :

$$\lim_{n \to \infty} a_n = a \tag{1}$$

En considérant les coefficients de z^2 il vient de même :

$$\lim_{n \to \infty} \sum_{h<k} p_{hn} \, p_{kn} = \frac{a^2}{2}$$

mais, comme $a_n^2 = \sum_{k=1}^{n} p_{kn}^2 + 2 \sum_{h<k} p_{hn} \, p_{kn} \longrightarrow a^2$, il en résulte que $\sum_{k=1}^{n} p_{kn}^2 \longrightarrow 0$; d'autre part $\alpha_n^2 \leqslant \sum_{k=1}^{n} p_{kn}^2$, d'où également :

$$\alpha_n \longrightarrow 0 \tag{2}$$

On voit ainsi que les conditions (1) et (2) sont nécessairement vérifiées. Notons que c'est ce théorème réciproque qui justifie pleinement la détermination de "loi des petites probabilités" que l'on donne souvent à la loi de Poisson.

RELATIONS ENTRE LES DISTANCES DE PAUL LEVY ET DE KY-FAN

N. ROBY

Soit $\mathfrak{X} = \mathfrak{X}(\Omega, B, \mathfrak{P})$ l'ensemble des variables aléatoires réelles définies sur un ensemble Ω probabilisé par la donnée d'une tribu borélienne B et d'une mesure de probabilité $\mathfrak{P}$.

Dans $\mathfrak{X}$, on définit, entre autres, deux convergences : la convergence en probabilité et la convergence en loi.

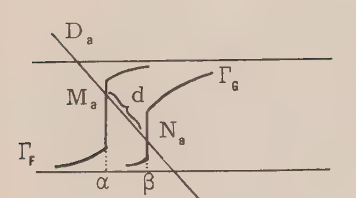
Il est connu que la convergence en probabilité est compatible avec une topologie métrisable, qu'on peut définir en munissant $\mathfrak{X}$ d'une structure d'espace vectoriel distancié. La distance de X à Y est la distance de X - Y à O, et la distance de X à O est définie par :

$$||X||_{K.F.} = \inf_{\varepsilon > 0} (\varepsilon \mid \mathfrak{P}(|X| > \varepsilon) < \varepsilon) . \quad \text{cf. (I)}$$

Nous appellerons cette distance la distance de Ky-Fan.

La convergence en loi, pour sa part, n'est pas compatible avec la structure d'espace vectoriel de $\mathfrak{X}$. Elle n'est, non plus, compatible avec aucune topologie séparée. Cependant, elle peut être définie par une distance que l'on met dans l'ensemble des fonctions de répartition ; par abus de langage, la distance des fonctions de répartition de X et Y s'appellera la distance de X et Y et, plus précisément, la distance de Paul Lévy. Elle est définie de la façon suivante. Soient X et Y $\in \mathfrak{X}$, F(x) et G(x) leurs fonctions de répartition respectives. Dans un plan orthonormé, soit Γ_F l'ensemble des points (x, y) qui vérifient la relation : $F(x) \leqslant y \leqslant F(x+o)$ (Γ_F est le graphe de F, rendu connexe par l'adjonction de segments verticaux aux points de discontinuité). Pour tout nombre a, la droite D d'équation y + x = a rencontre Γ_F en un point M_a et Γ_G en un point N_a. La distance de Paul Lévy est alors définie par :

$$d_L(X, Y) = \sup_{a \in R} M_a N_a .$$



Il est connu qu'entre les deux distances ainsi définies on a la relation :

(a) $\boxed{d_L(X,Y) \leqslant \sqrt{2}\,||X-Y||_{K.F.}}$ (cf. (II), p.51).

Nous voulons ici montrer d'autres relations entre ces distances, mais pour donner une unité à cet article nous allons redémontrer (a). Il suffit de montrer que, pour tout a, on a : $M_aN_a \leqslant \sqrt{2}\,||X-Y||_{K.F.}$.

Posons $M_aN_a = d$. Soient α et β les abscisses respectives de M_a et N_a. La relation à démontrer étant symétrique en X et Y, on peut supposer $\alpha \leqslant \beta$. Alors, $\beta - \alpha = \frac{d}{\sqrt{2}}$. Il est évident que l'on a :

$$F(\alpha+o) - G(\beta) \geqslant \frac{d}{\sqrt{2}}$$

c'est-à-dire :

$$\mathcal{P}(X \leqslant \alpha) - \mathcal{P}(Y < \beta) \geqslant \frac{d}{\sqrt{2}}$$

Or, pour deux évènements A et B, il est évident que :

$$\mathcal{P}(A) - \mathcal{P}(B) \leqslant \mathcal{P}(A \cap \overline{B})$$

Donc aussi :

$$\mathcal{P}(X \leqslant \alpha \cap Y \geqslant \beta) \geqslant \frac{d}{\sqrt{2}}$$

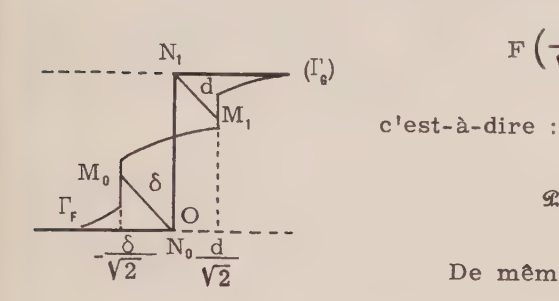
Mais l'évènement $(X \leqslant \alpha) \cap (Y \geqslant \beta)$ est inclus dans l'évènement : $|Y-X| \geqslant \frac{d}{\sqrt{2}}$. Donc, à fortiori :

$$\mathcal{P}\left(|Y-X| \geqslant \frac{d}{\sqrt{2}}\right) \geqslant \frac{d}{\sqrt{2}}$$

On conclut de là sans peine que $\frac{d}{\sqrt{2}} \leqslant ||Y-X||_{K.F.}$ c.q.f.d.

De (a) résulte le résultat bien connu, que la convergence en probabilité entraîne la convergence en loi.

Considérons maintenant le cas particulier où Y est égal à 0 avec la probabilité 1. Si l'on coupe Γ_F et Γ_G par les droites D_0 et D_1, on obtient une situation schématisée par la figure ci-contre ; on pose $M_1N_1 = d$, $M_0N_0 = \delta$.



On a visiblement :

$$F\left(\frac{d}{\sqrt{2}} + O\right) \geqslant 1 - \frac{d}{\sqrt{2}}$$

c'est-à-dire :

$$\mathfrak{P}\left(X > \frac{d}{\sqrt{2}}\right) \leqslant \frac{d}{\sqrt{2}}$$

De même :

$$\mathfrak{P}\left(X < -\frac{\delta}{\sqrt{2}}\right) \leqslant \frac{\delta}{\sqrt{2}}$$

Comme l'évènement $|X| > \frac{d+\delta}{\sqrt{2}}$ est inclus dans l'évènement $\left(X > \frac{d}{\sqrt{2}}\right) \cup \left(X < \frac{-\delta}{\sqrt{2}}\right)$ on a donc :

$$\mathfrak{P}\left(|X| > \frac{d+\delta}{\sqrt{2}}\right) \leqslant \mathfrak{P}\left(X > \frac{d}{\sqrt{2}}\right) + \mathfrak{P}\left(X < -\frac{\delta}{\sqrt{2}}\right) \leqslant \frac{d+\delta}{\sqrt{2}}.$$

Or, pour tout $\varepsilon > \frac{d+\delta}{\sqrt{2}}$ on vérifie immédiatement que $\mathfrak{P}(|X| > \varepsilon) < \varepsilon$, par conséquent :

$$||X||_{K.F.} \leqslant \varepsilon \quad \forall\, \varepsilon \geqslant \frac{d+\delta}{2}$$

Finalement :

$$||X||_{K.F.} \leqslant \frac{d+\delta}{\sqrt{2}} \leqslant \sqrt{2}\; d_L(X,O)$$

Ainsi, en tenant compte de la formule (a), on a :

$$\text{(b)} \quad \boxed{d_L(X,O) \leqslant \sqrt{2}\,||X||_{K.F.} \leqslant 2\, d_L(X,O)}$$

En ajoutant à chacune des variables en présence une constante a (ce qui ne change aucune des deux distances), et changeant X en X-a, on a la formule générale suivante :

$$\text{(b')} \qquad \boxed{d_L(X,a) \leqslant \sqrt{2}\, \|X-a\|_{K.F.} \leqslant 2\, d_L(X,a)}$$

Elle montre que la convergence en loi et la convergence en probabilité vers une limite presque-sûrement constante sont deux phénomènes équivalents, ce qui est un résultat de Cantelli (cf. (I), et (III) p. 168).

Supposons maintenant, en outre, que dans la relation (b) X soit presque-sûrement toujours $\geqslant 0$ ou toujours $\leqslant 0$. Alors, l'un des deux nombres d ou δ est nul. Il est donc inutile de le majorer par $d_L(X,O)$, ce qui a pour effet de diviser par deux le troisième membre de (b). On obtient donc un résultat qui, sous sa forme généralisée, s'écrit :

$$\text{(c)} \qquad \boxed{d_L(X,a) = \sqrt{2}\, \|X-a\|_{K.F.}} \quad \underline{\text{si, p. s., } X \geqslant a \text{ ou } X \leqslant a.}$$

On en déduit, par exemple, une relation où ne figure que la distance de Paul Lévy. Quels que soient X et Y $\in \mathfrak{X}$ on a en effet :

$$d_L(X,Y) \leqslant \sqrt{2}\, \|X-Y\|_{K.F.} \leqslant 2\, d_L(X-Y,\, O)$$

c'est-à-dire :

$$\text{(d)} \qquad \boxed{d_L(X,Y) \leqslant d_L(|X-Y|,\, O) \leqslant 2\, d_L(X-Y, O)}$$

REFERENCES

(I) - M. FRECHET - Généralités sur les probabilités. Premier livre.

(II) - P. LEVY - Théorie de l'addition des variables aléatoires (1954).

(III) - LOEVE - Probability Theory.

NOTE SUR LA DISTANCE D'UNE VARIABLE ALÉATOIRE X A L'ORIGINE

Edouard BONET

a) Définition de M. Ky-Fan.

La distance d(O,X) est définie comme la borne inférieure des des valeurs x positives telles que :

$$\text{Prob}\,[\,|X| > x\,] < x$$

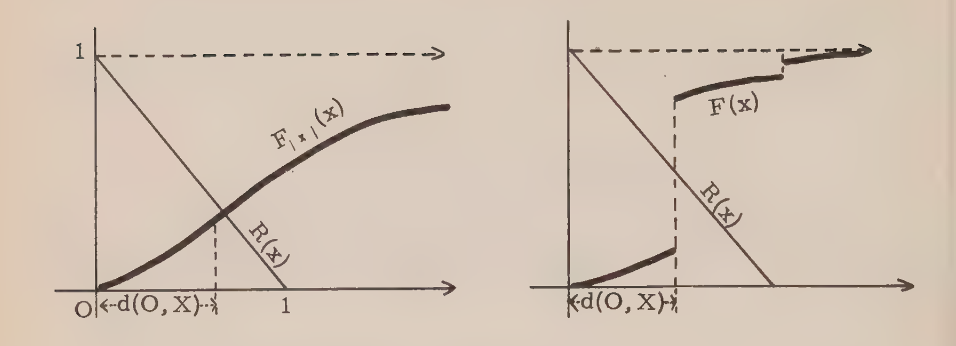
Il faut noter que cette inégalité n'impose aucune condition quand x est supérieur à 1, parce que toute probabilité est toujours inférieure ou égale à 1. Cela entraîne quelques conséquences :

1/ - $d(O,X) \leqslant 1$;

2/ - $d(O,X) = 1 \Longleftrightarrow \text{Prob}\,[\,|X| < 1] = 0$;

3/ - Quand X est une variable certaine comprise entre -1 et 1 alors cette distance devient la valeur absolue de X, mais cette propriété n'est plus vraie quand X a une valeur plus grande et en conséquence on ne peut pas considérer cette distance comme une extension du concept de distance entre variables certaines.

On obtient une représentation graphique de la façon suivante : si F(x) est la fonction de répartition de $|X|$ et R(x) la droite menée par les points (0,1), (1,0) alors d(O,X) est la borne inférieure des valeurs positives de x pour lesquelles $F(x) > R(x)$.



b) Proposition de distance.

On peut proposer comme définition de d(O,X) la borne inférieure des valeurs positives de x pour lesquelles :

$$\text{Prob}\ [|X| > x] < x - [x] \quad \text{avec } [x] = \text{partie entière de } x$$

Cette nouvelle définition entraîne :

1/ - $d(O,X) < \infty$;

2/ - Quand X est une variable certaine alors d(O,X) devient la valeur absolue de cette variable.

Cette distance est une extension de la distance entre variables certaines.

Représentation graphique.

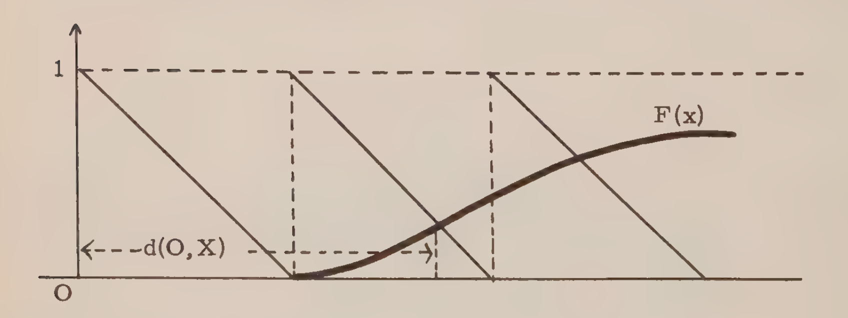


TABLE GÉNÉRALE DES MATIÈRES
VOLUME IX – 1960

Pages

IMP. LOUIS-JEAN — GAP
Dépôt légal n° 146 — 1961